The University of Southern Mississippi

The Aquila Digital Community

Dissertations

Summer 8-1-2021

Test-Driving Interventions for Teachers: A Proactive Method for Improving Treatment Integrity

Stefanie Schrieber

TEST-DRIVING INTERVENTIONS FOR TEACHERS: A PROACTIVE METHOD

FOR IMPROVING TREATMENT INTEGRITY

by

Stefanie R. Schrieber

A Dissertation
Submitted to the Graduate School,
the College of Education and Human Sciences
and the School of Psychology
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved by:

Dr. Brad Dufrene, Committee Chair
Dr. Evan Dart
Dr. Joe Olmi
Dr. Sara Jordan

August 2021

Published by the Graduate School

THE UNIVERSITY OF
SOUTHERN
MISSISSIPPI.

ABSTRACT

Test-driving is an antecedent method utilized to improve the extent to which teachers adhere to implementation procedures of student interventions (i.e., treatment integrity). The current study aimed to extend the literature on the test-driving intervention with three teacher-student dyads in a high school setting using a multiple baseline design. Teacher/student dyads were recruited based on teacher referral to the schools' behavioral consultants. Recruited teachers began in a consultation-as-usual phase, where procedures were implemented consistent with the problem-solving consultation model (e.g., problem identification, problem analysis, treatment implementation, and treatment evaluation). Teachers who demonstrated 50% or lower adherence to treatment components were then recommended to participate in the test-driving intervention. During the test-drive, teachers were trained on four different student interventions using behavioral skills training. Teachers were then expected to implement these four different interventions with 100% integrity across two days. Following this test-drive of interventions, teachers rated the acceptability and feasibility of the interventions on the URP-IR and independent of these ratings, selected their most-preferred intervention from those that were implemented with 100% integrity. Researchers expected that TI may improve if teachers were able to select an intervention to implement after they were able align their expectations of each intervention to their actual performance. Results were primarily analyzed via visual analysis which indicated little to no changes in TI data nor student behavior data. Additional research is needed to determine the relationship between choice, acceptability, and treatment integrity.

ACKNOWLEDGMENTS

DEDICATION

I'd like to dedicate this work to all of those that have supported me during my graduate school journey. A special thank you goes out to my mother, who has always encouraged me to follow my passions and recognizes my goal to be a life-long student. I'd also like to thank my fiancé Eric, who's been my fiancé for three years now, but never complains about the distance or time between us. I have to also give a shout out to Dr. Chad Benis, who can turn around a draft faster than you can say "Chad Benis." Other shout outs to my Skull friends, who are still dealing with my limited bank account funds, Squiddies, who are the best parents and friends I could ask for, and to my new Boston friends that have made internship year one I'll never forget!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

TI   Treatment Integrity

AEB   Academically Engaged Behavior

CHAPTER I - INTRODUCTION

Problem-solving consultation is an indirect service delivery framework utilized to address client behavior (Bergan & Kratochwill, 1990). Roles within the problem-solving consultation framework include a consultant, consultee, and client. The consultee's role is to communicate concerns about a client to the consultant. The consultant serves as the expert and provides the consultee with strategies to address the communicated concerns. The consultee is then responsible for implementing the recommended strategies with the client. The overarching goal of problem-solving consultation is for the consultant to provide the consultee with necessary skills to change the client's behavior.

Problem-solving consultation is the dominant model of service delivery for school psychologists (Erchul & Martens, 2012). Through problem-solving consultation, school psychologists are able to provide a variety of indirect academic, behavioral, and mental health services (Gutkin, 1996; Kratochwill & Stoiber, 2000). Compared to direct service methods, where the school psychologist is responsible for the implementation of interventions with clients, indirect methods capitalize on already limited school resources. The indirect approach of problem-solving consultation allows school psychologists to serve a larger population of students by providing teachers with strategies that can promote behavior change across multiple students (e.g., the Good Behavior Game; Barrish, Saunders, & Wolf, 1969) or be generalized to address similar student problem behaviors in the future (Erchul & Martens, 2012).

The traditional problem-solving consultation framework follows a four-stage problem solving process (Bergan & Kratochwill, 1990). This process includes (a) problem identification, (b) problem analysis (c) treatment implementation, and (d)

treatment evaluation. Problem identification is the process by which the consultant (i.e., the school psychologist) and the consultee (i.e., the teacher) meet to discuss the behavior the teacher would like to change in the client (i.e., the student). During this stage, the consultant and consultee develop definitions for the target behavior, discuss antecedents and consequences that may be maintaining the target behavior, and develop methods for collecting baseline data on the behavior. During the second stage, problem analysis, the school psychologist and teacher meet to evaluate the baseline data and develop an intervention to address the behavior. Next, the teacher implements the selected intervention, continues to collect data on the target behavior, and briefly meets with the school psychologist to discuss any challenges or issues related to implementation. The final stage in the problem-solving process is treatment evaluation. During this stage, the school psychologist and teacher review all student behavior data and treatment integrity data if available to determine whether the intervention was effective for promoting change in the client's target behavior or if additional training is warranted.

Intervention effectiveness, as described in the traditional problem-solving consultation framework, is dependent on the evaluation of student progress. Thus, decisions regarding the effectiveness of an intervention are often made without considering whether the treatment was accurately implemented by the consultee. This raises concerns, as student response to treatment has been shown to be related to accurate implementation of interventions (i.e., greater improvements in student behavior are observed when interventions are implemented as they are intended; Noell, Gresham, & Gansle, 2002). Because accurate implementation of treatment is critical to improve

student outcomes, intervention implementation should be assessed in conjunction with student progress prior to making decisions about intervention effectiveness.

Treatment Integrity

The degree to which an intervention is implemented as it was intended is known as treatment integrity (TI; Gresham, 1989). TI provides data to support whether the independent variable (i.e., the selected treatment) was systematically and accurately implemented. It is a key reason why researchers can conclude the presence of functional relationships between independent and dependent variables (Shadish, Cook, & Campubell, 2002). Noell, Gresham, and Gansle (2002) demonstrated that student outcomes differed in response to varying levels of TI on a computer administered math intervention. When the intervention was implemented with 33% and 67% accuracy, students demonstrated little to no improvements in the number of digits correct when compared to baseline; however, when the intervention was implemented with 100% accuracy, all students demonstrated increasing levels of digits correct per minute. Thus, decisions made without the consideration of both TI and student outcomes are subject to error, as student non-response to treatment may be attributed to poor implementation of the intervention (Gresham, 1989; Noell, Gresham & Gansle, 2002).

Traditionally, the evaluation of TI focused primarily on the adherence to intervention procedures as prescribed (Peterson, Homer, & Wonderlich, 1982). Adherence is measured by listing the components of the selected intervention and determining whether each component was implemented. Although more recent evaluations of TI have suggested that there may be dimensions that should be evaluated in addition to adherence (e.g., dosage of treatment, and quality of implementation; Sanetti

and Kratochwill, 2009); these dimensions have not been agreed upon in the literature. Additionally, one may argue that without adherence to treatment procedures, other dimensions of TI cannot be accurately assessed. For example, if an intervention is only implemented with 50% adherence to treatment components, then measurements of dosage or quality may not be great indicators of TI as they would not capture whether the intervention was implemented in its entirety. Thus, the assessment of TI will be primarily based on the level of adherence to treatment procedures.

There are various methods that may be utilized to assess adherence to treatment components. Three primary methods of TI assessment include systematic direct observation (SDO), self-report, and permanent product (Lane, Bocian, MacMillan & Gresham, 2004). SDO requires an individual other than the consultee to watch and record the consultee's implementation of the intervention (Lane et al., 2004). Typically, the observer uses a checklist of intervention components to record which components were and were not implemented by the interventionist. The number of correct steps is then divided by the number of total possible steps and converted to a percentage (Hagermoser Sanetti, & Kratochwill, 2005). Although one may suggest that teachers may react to the presence of an observer, Codding, Livanis, Pace and Vaca (2008) evaluated teacher reactivity to the presence and absence of observers and demonstrated that the presence of an observer produced no differences in the level of TI. Therefore, the primary limitation of SDO is the intensity of resources required (e.g., observers, time). A more indirect measure of TI is self-report, which typically requires the teacher to complete a checklist of intervention components independently (Hangermoser et al., 2005). Although self-report is more resource efficient than systematic direct observation, research has shown

that individuals often overestimate their adherence to the intervention (Gresham et al., 2000). Permanent products may also be used to measure the integrity of an intervention (Sanetti & Kratochwill, 2009). Data collected via permanent product are recorded after an intervention occurs using tangible objects or environmental changes as the measure of TI. An example of a permanent product is a completed worksheet from an academic intervention. This method is resource efficient, as it does not require real-time observations or additional school personnel to measure; however, there are often components of interventions that do not leave permanent products behind that can be used to determine adherence to procedures (e.g., statements of behavior specific praise; Sanetti & Collier-Meek, 2013). Gresham, Dart, and Collins (2017) evaluated the reliability of these three methods of TI to determine whether permanent product and self-report measures produced similar results to SDO. These methods were evaluated based on teachers' implementation of the good behavior game. Although self-report and permanent product measures were dependable, they required a larger amount of data to accurately assess TI in comparison to SDO. Additionally, TI collected from permanent product may not capture all necessary components, especially when compared to SDO. for accurate. A second finding of this study was that the majority of variance in the level of TI was due to differences in teachers (i.e., implementers).

Although methods and dimensions of TI have received increased attention over the years, it is still underreported in the literature. A systematic review of school psychology journals between 1995 and 2008 indicated that 37.2% of articles did not include quantitative integrity data nor mentioned monitoring TI data (Sanetti, Dobey, & Gallucci, 2014). Additionally, 29.1% did not include a definition or a reference to a

definition of the independent variable. In total, 29.1% of the studies were labeled as "high risk" for issues related to TI, meaning they did not provide quantitative TI data, did not mention assessment of TI, and were at risk for inaccurate implementation. Thus, there are many articles that were published regarding the effectiveness of interventions without consideration of TI data. Of the TI data that was reported, the average level of integrity was 93.6%. Although this statistic indicates a high level of TI, it should be interpreted with caution. First, methodologies for assessing TI were not reported, such as the type of individual implementing interventions (e.g., experienced researcher vs. teacher), and assessment procedures (e.g., self-report vs. direct observation). Second, this statistic may not be representative of TI in more applied settings outside of the research context. Third, there is a potential for publication bias, as studies with poor TI may be less likely to be published than studies that do not report any data related to TI.

Difficulties implementing interventions as planned have been consistently observed in the literature (Noell et al., 2005; Sanetti & Kratochwill, 2009) indicating a need for methods to improve consultees' TI. Research within the problem-solving consultation literature has shown that although teachers initially demonstrate high levels of TI, within ten days after training, levels of integrity often decrease substantially ranging from 0-65% adherence to intervention procedures (Mortenson & Witt, 1998; Noell, Witt, Gilbertson, Rainer, & Freeland, 1997; Sanetti & Kratochwill, 2008). Additionally, Wickstrom, Jones, LaFleur and Witt (1998) evaluated TI within the traditional problem-solving consultation framework and found low levels of TI. In this study, TI was assessed through teacher completion of a Baseline and Intervention Record Form (BIRF), assessment of stimulus product use, and treatment use. The BIRF was a

data collection form completed by teachers and was used to code student behaviors every day of the week at specified intervals. Stimulus product use was assessed based on the presence and visibility of intervention materials to the target student. Treatment use referred to the teacher's actual implementation of the intervention following the student's engagement in the target behavior. Twenty-eight teachers that had referred students for problem behaviors participated in the study. Teachers were instructed on intervention procedures using the traditional problem-solving consultation framework as described by Bergan and Kratochwill (1990). Results indicated that teachers' average integrity of the BIRF completion was 54% (range = 0-100%), average stimulus product use was 62% (range = 0-100%), and average treatment use following presence of the target behavior was 4% (range = 0-21%). Based on the consistent demonstrations of low TI for teacher-implemented interventions and the importance of TI in treatment effectiveness, identification of procedures that may increase and/or maintain high levels of TI is necessary.

One way that TI is being addressed is through implementation science. Implementation science is utilized across many different areas of research outside of school psychology, including the medical and healthcare fields (Rogers, 2003). Broadly, implementation science is defined as the process of identifying methodologies and processes to promote the adoption of evidence-based interventions in applied contexts (McHugh & Barlow, 2010). When applied specifically to the field of school psychology, implementation science is the process of identifying barriers to implementing evidenced-based practices in the schools, improving TI of interventions being implemented by school personnel, and identification and implementation of culturally diverse and

appropriate interventions in the school (Forman et al., 2013). For the purposes of this study, emphasis will be placed on the methodologies to improve TI of interventions implemented by school personnel.

<div align="center">Strategies to Promote TI</div>

Strategies to promote TI have primarily relied on responsive methods, meaning that many strategies are not utilized until after low levels of TI have been observed. Procedures that have been largely utilized as responsive methods in the literature include implementation planning, treatment integrity planning protocol (TIPP), performance feedback, self-monitoring, participant modeling and role play, coaching (Reinke, Stormont, Herman, & Newcomer, 2014) and motivational interviewing (Noell, Witt, Gilbertson, Ranier, & Freeland, 1997; Sanetti, Collier-Meek, Long, Byron, & Kratochwill, 2015; Sanetti & Kratochwill, 2009; Sanetti & Collier-Meek, 2019). Preventative methods, or methods that decrease the likelihood that low integrity ever occurs, are mostly absent in the literature, except for test-driving (Dart, Cook, Gresham & Chenier, 2012). Although the literature has largely evaluated coaching and performance feedback as responsive interventions, there is some evidence to suggest that they may also be used as a preventative method. The dearth of research evaluating preventative methods is concerning, as consequent approaches may be considered "wait to fail" models. "Wait to fail" models, previously used in the identification of students for special education, delay intervention supports until performance is substantially behind what is expected (Otaiba, Wagner, & Miller, 2014). The use of consequent methods to improve TI are similar to "wait to fail" models, as they allow the teacher to fail at

accurate implementation prior to providing any supports. The delay in supports to address low integrity can subsequently delay improvements in student outcomes.

Responsive Methods

Responsive methods are used to respond to low treatment integrity. A consultant following the traditional problem-solving consultation framework may recognize non-adherence to treatment components through brief meetings during the treatment implementation stage. The consultant may then determine a method to improve TI of the intervention. Methods utilized to improve integrity after low levels of TI have already been demonstrated include performance feedback (Noell et al., 1997), self-monitoring (Simonsen, MacSuga, Fallon, & Suagi, 2013), participant modeling and role play (Sanetti & Collier-Meek, 2015), motivational interviewing (Miller & Rollnick, 2002), coaching (Reinke et al., 2014), implementation planning (Sanetti et al., 2015), and treatment integrity planning protocol (Sanetti & Kratochwill, 2009).

Performance Feedback. The most common strategy with the most evidence supporting its effectiveness for promoting TI is the use of performance feedback (Noell et al., 1997). Within this method, direct observations of TI are conducted by the consultant. The consultant and consultee then meet to discuss the implementation of the intervention. During these meetings, the consultant reviews both TI and student outcomes with the teacher and provides specific feedback about the correct, incorrect, or missed steps during implementation (Noell, 2010; Sanetti, Fallon & Collier-Meek, 2011). Performance feedback can be conducted a variety of different ways and may incorporate components such as negative reinforcement and self-monitoring procedures (DiGennaro, Martens, & Kleinmann, 2007; Noell et al., 1997, 2005; Simonsen, Meyers & DeLuca, 2010). Often,

negative reinforcement is used to allow teachers to escape meetings with the consultant contingent on high levels of integrity. Self-monitoring procedures require the teacher, in addition to the consultant, track their integrity during implementation. Consultees then meet to compare their ratings of TI to the consultant's ratings. In general, performance feedback can be provided on a daily schedule, weekly schedule, or contingent on the level of TI. Feedback can also be provided through multiple different outlets, including face-to-face meetings, email, or on paper (Fallon, Collier-Meek, Maggin, Sanetti, & Johnson, 2015). In a meta-analysis, performance feedback was moderately effective for improving TI for both experimental (IRD = .64, r = .71) and quasi experimental studies (IRD = .52, r = .70; Solomon, Klein & Politylo, 2012). Performance feedback also requires additional resources that may not be feasible given already limited resources in the school. For example, an observer needs to be present at the time of the intervention and additional meetings need to be held following intervention sessions to provide feedback. Overall, performance feedback has been identified as an evidence-based intervention; however, it there is also evidence to suggest even greater improvements in TI when used in combination with other strategies such as self-monitoring (Sanetti & Collier-Meek, 2019).

  Self-Monitoring. TI has also been improved via self-monitoring procedures alone (Simonsen, MacSuga, Fallon, & Suagi, 2013). Self-monitoring procedures require teachers to track the frequency or adherence to the intervention components while the consultant conducts their own integrity check. The teacher and consultant will then meet to compare the levels of TI and discuss any discrepancies. In a study conducted by Simonsen and colleagues (2013), five teachers were trained to increase their rate of

behavior specific praise (BSP) statements. An alternating treatments design was used to compare different forms of self-monitoring including a basic tally count, a frequency count using a golf counter, and estimations of the rate of BSP statements made per minute. SDO was used to compare to the teachers self-monitoring procedures and determine the level of accuracy. Results indicated that accuracy of BSP statements varied across all methods of self-monitoring. Additionally, even after the best treatment phase was implemented, three out of five teachers entered a performance feedback phase due to decreases or low levels to the treatment (i.e., BSP statements). Although self-monitoring procedures have been an effective strategy for promoting behavior change, evidence suggests it is best used in combination with other strategies (Sanetti & Collier-Meek, 2019). Therefore, self-monitoring procedures may require even more resources than other consequent methods because they are not as effective when used independently.

Participant Modeling and Role Play. Participant modeling and role-playing are two similar strategies often used as responsive methods to improve integrity data (Sanetti & Collier-Meek, 2015). Participant modeling is conducted in vivo, and role-playing is conducted outside of the implementation setting. In participant modeling, the consultant first demonstrates the intervention with the client(s). Then, consultees are provided the opportunity to practice with the client and receive real-time feedback from the consultant. During role-play methods, the consultant first demonstrates the intervention with the consultee as the client, and then the consultee practices with the consultant as the client. During role play demonstrations, potential barriers to implementation are identified and often included to provide the consultee strategies to overcome them. Participant modeling has been demonstrated as an effective intervention when used for teachers exhibiting low

11

to moderate levels of TI who have already undergone direct training and implementation planning (Sanetti & Collier-Meek, 2015). Role-playing has also been demonstrated as effective for improving teacher fidelity, especially when used in combination with other strategies such as direct training (Sterling-Turner, Watson, Wildmon, Watkins & Little, 2001).

Motivational Interviewing. Motivational interviewing is another strategy that may promote TI. Motivational interviewing was originally developed as a method for treating concerns related to substance abuse. More recently, it been adapted for use in school-based consultation. Motivational interviewing may be best used when consultants have reason to believe teachers can implement the intervention with high integrity but are unable to maintain these levels across time (e.g., a performance deficit). During motivational interviewing, the consultant follows the consultees lead, and carefully guides the conversation to assist with the identification of reasons for change. The consultant can then use this information to help develop a plan to support the identified changes. Although motivational interviewing has been successful in promoting the use of EBPs in school, community, and mental health centers (Hettema, Ernst, Williams & Miller, 2014), it does require a skilled individual to lead the process, and schools may not have access to such personnel.

Implementation Planning. Implementation Planning, adapted from the health action process approach (HAPA), has also been utilized to promote TI (Sanetti et al., 2015). The two main strategies within implementation planning are action planning and coping. Action planning is the process of reviewing and discussing all intervention steps, revising intervention steps based on the needs of the student and the teacher, and

recording answers to teacher's questions about when the intervention will occur, dosage, and identification of additional resources needed to implement the intervention. Coping is the process of identifying and addressing potential barriers of implementation. Sanetti and colleagues (2015) demonstrated improvements in adherence to treatment components following implementation planning. In their study, when adherence percentages were 80% or below for a minimum of two days, an implementation planning meeting occurred. During this meeting, the consultant and consultee discussed the action plan and coping methods. Although implementation planning resulted in improvements in adherence to treatment procedures, decreasing trends in integrity were observed at 1-month and 2-month follow ups.

Treatment Integrity Planning Protocol. Another method that may be used to promote TI is the use of the Treatment Integrity Planning Protocol (TIPP; Sanetti & Kratochwill, 2009). TIPP is a strategy that can be implemented within the problem-solving consultation problem solving process. During the problem analysis stage of the process, a meeting is arranged for the teacher and school psychologist to operationally define the steps of the intervention and identify the interventionist, location, and dosage of implementation. Next, multiple dimensions of TI including interventionist behaviors, student behaviors, and dosage are determined and divided into steps for the integrity form. The method for assessing the steps (e.g., self-report, direct observation), response format for each step (e.g., checklist, Likert scale), and a remediation plan for poor integrity are developed during the meeting. Finally, the assessment method is created and a schedule for frequency of assessment is determined. TIPP has been effective for increasing levels of TI and accuracy of teacher self-report measures of integrity (Sanetti

13

& Kratcochwill, 2009). Additional evaluations of TIPP by Sanetti and Kratochwill (2011) have only resulted in moderate and more stabilized improvements, as compared to variable levels, of TI. Although TIPP utilizes both antecedent and consequent methods, evaluations of this strategy are limited, have relied on teacher self-report and permanent product assessment methods, and have not resulted in high levels of TI.

Coaching. This method involves an ongoing 'coach' or expert in the field that is available to assist with implementation of an intervention. Coaching serves to bridge the gap between an expert's didactic presentation and application of the intervention in the classroom setting. There is evidence to suggest that when teachers are provided with access to additional supports, they are more likely to implement an intervention (Driscoll et al., 2011). Although coaching procedures can vary by individual, it often includes components of other evidence-based procedures (e.g., performance feedback, modeling action planning, and reviewing).

Although these methods have demonstrated improvements in teacher integrity, they have largely been evaluated as responsive strategies. Ideally, methodology for promoting integrity should be proactive and prevent lapses in integrity before implementation ever occurs. Currently, there may only be one method that has been identified to prevent individuals from implementing interventions with low integrity.

Preventative Methods

Recall that preventative methods to promote TI include strategies that serve to prevent lapses of integrity from ever occurring. Preventive methods use proactive strategies that provide teachers supports to promote high integrity before implementation rather than using responsive or "wait to fail" methods after low integrity is observed.

Although some of the methods described in the responsive methods section may be utilized as a preventative method, evidence in the literature has only supported them as responsive methods when low TI has already been observed. Test-driving may be a preventative method for promoting high levels of TI in the literature.

Test-Driving Interventions

Test-driving is a preventative strategy that allows teachers to implement, or test, multiple different evidence-based interventions and choose the one they like best to implement over a longer period of time. In a single study, Dart, Cook, Collins, Gresham & Chenier (2012) indicated that teachers may be more inclined to follow through with intervention procedures if they have the opportunity to test and choose an intervention procedure. Although teachers may perceive an intervention as acceptable prior to implementation, an acceptability rating cannot be given until teachers experience the actual performance of the product. Therefore, providing teachers an opportunity to implement each intervention will allow teachers to adjust their acceptability of each intervention. It was hypothesized by Dart and colleagues (2012) that teachers would choose the intervention that most matched their expectations (i.e., had the greatest acceptability).

Dart and colleagues (2012) conducted the first empirical evaluation of test-driving interventions with teachers to improve TI and student outcomes. Initially, six kindergarten to fourth grade student-teacher dyads were recruited for the study; however, two dyads were excluded due to high adherence to treatment procedures during baseline. Four evidence-based interventions were pre-selected by the researchers based on a

similar number of required resources. The selected interventions included self-monitoring, modified Check-in/Check-out, response cost, and behavior specific praise.

Results were evaluated utilizing a single-case, multiple baseline design across two linked teachers. Instead of a baseline condition, teachers started in a consultation-as-usual condition (i.e., the consultant chose an evidence-based intervention from the four possible options to address the problem behavior). Then, the first teacher in each link 'test-drove' the remaining three interventions. The test drive served as a brief experimental analysis, in which teachers briefly tested each of the four intervention procedures to observe how they were implemented and their impact on student outcomes. Teachers then ranked the interventions from most to least favorite and were asked to implement the procedure they ranked the highest. The second teacher in each linked pair then implemented the intervention that was selected by the first teacher. This was to confirm that it was the teacher's ability to choose an intervention rather than the intervention itself that resulted in improvements in integrity. Finally, the second teacher in each linked pair experienced the test-driving procedure with the remaining two interventions.

All teachers demonstrated low or variable levels of TI during the consultation-as-usual condition. After the test drive, the first teacher in each linked pair demonstrated increases in TI to 100% following their choice to implement the preferred intervention. For the second teacher in each linked pair, they continued to demonstrate low levels throughout the consultation-as-usual condition and during implementation of the first teacher's chosen intervention. After implementation of the first teacher's best perceived intervention, they experienced the test drive condition. Following test-driving, these teachers also demonstrated improvements in TI during the preferred intervention

16

condition. Additionally, improvements in student behavior were observed during the preferred intervention condition for three out of four teachers. Although the test-driving intervention was successful in promoting TI and student outcomes, the results should be considered in light of several limitations.

First, the researchers failed to provide data to support the implementation of the test drive condition. As noted above, TI data are necessary to accurately evaluate the effectiveness of an intervention. Overall, the lack of data provided during the test drive condition (i.e., student outcomes and procedural integrity) limit the internal and external validity of the study. Therefore, no statements can be made regarding the functional relationship between TI and test-driving. Second, the study did not meet What Works Clearinghouse (WWC; Kratochwill et al., 2010) single-case design standards, in that some phases included only three data points and only two replications of experimental effect were present across the linked teachers. Finally, the researchers did not collect maintenance data, limiting decisions regarding the effectiveness of the intervention across time.

Despite these limitations, the test-driving intervention is a potential valuable antecedent method for improving TI that needs further evaluation to determine its effectiveness. There are two major components that may contribute to the success of test-driving as an intervention method. These include the incorporation of treatment acceptability and teacher choice. These two factors of test-driving provide opportunities for teachers to choose an intervention that aligns with their expectations which may result in improvements to adherence to treatment components.

Treatment Acceptability

Treatment acceptability can be defined as the degree to which an intervention is perceived as fair, appropriate, reasonable, intrusive, and consistent with expectations of what the treatment should do (Kazdin, 1980). Similar to treatment integrity, researchers have identified many other factors of acceptability including: problem severity, treatment approach, side effects of treatment, time needed to implement, and cost (Miltenberger, 1990). Acceptability is often a component measured within assessments of social validity. For example, acceptability is a factor measured within the Usage Rating Profile – Intervention Revised (URP-IR), which is a rating scale that can be completed by teachers to measure the social validity of interventions (Neugebauer, Chafouleas, Coyne, McCoach, & Briesch, 2016). Examples of items on the URP-IR within the acceptability factor include, "This intervention is an effective choice for addressing a variety of problems," "The intervention is a fair way to handle the child's problem behavior," and "I would have positive attitudes about implementing this intervention." Satisfaction is another component often included within social validity. Similar to consumer satisfaction, a term borrowed from industrial/organizational psychology, once an individual uses a product (or intervention), they evaluate whether its performance meets their expectations. If the actual performance of the intervention does not meet their expectations, individuals are likely to be unsatisfied with the product and therefore rate the product as having low acceptability (Witt & Elliot, 1985). Subsequently, an individual may be less likely to use a product or implement an intervention as originally intended. Research evaluating the relationship between treatment acceptability and TI have not produced clear and consistent results.

Sterling-Turner and Watson (2002) examined the relationship between treatment acceptability and integrity within a sample of undergraduate students. The study required participants to read a case description with a treatment plan, undergo training to implement the intervention, and then implement the intervention with a client. Results indicated that there were no significant correlations between treatment acceptability and integrity. It is important to consider that Sterling-Turner and Watson (2002) utilized undergraduate participants, whose perception of acceptability and TI are likely not directly comparable to a teacher's acceptability and TI in a school setting. Therefore, these results should be considered lightly. Noell and colleagues (2005) conducted a consultation study with 45 teacher-student dyads. Although the study's primary research question was related to how different consultation methods affected TI, secondary analyses were conducted to determine correlations between TI and treatment acceptability. Although small, non-significant correlations were found, this was not the study's primary purpose, and these results are limited as correlational findings and do not indicate causal relationships. Allinder and Oats (1997) conducted a study assessing acceptability of interventions and identified a relationship between treatment acceptability and TI. In their study, teachers rated their acceptability for curriculum-based measurement (CBM) with students. Teachers who perceived CBM as more acceptable were found to utilize CBM more frequently. Additionally, for those teachers with higher acceptability, greater outcomes were observed for the students. This may indicate a functional relationship between high acceptability and high levels of TI, as high levels of TI have been demonstrated to improve student outcomes. Although research investigating this relationship is slim, the possibility of high acceptability being directly related to high

19

levels of TI should not be dismissed. Research suggests that teachers are more likely to be satisfied with an intervention when provided the opportunity to collaborate with the consultant (Wenger, 1979), therefore leading to increased levels of TI. Because overall satisfaction and acceptability of an intervention is based on discrepancy between one's perception of the performance and the actual performance, consumers cannot rate these factors without first implementing the intervention with integrity. When teachers can align these perceptions of expected and actual performance, it may assist with rating acceptability of the interventions. However, a rating from a survey/questionnaire that indicates high acceptability of an intervention may not necessarily coincide with which intervention a teacher would choose to implement long term. Thus, choice may be an important element to include that is a more dependable indicator of which intervention teachers may be more likely to implement with integrity in comparison to acceptability surveys alone.

<div align="center">Choice</div>

In the school-based literature, choice is often used as an antecedent intervention to reduce the likelihood that students engage in problem behavior; however, it has also been demonstrated as effective for increasing academic engagement and improving task performance across multiple studies (Bannerman et al., 1990). Choice can be incorporated into daily routines, including choice of reinforcers, interventions, academic tasks, order of tasks, type of task, and has largely been used with student populations. Tiger, Hanley, and Hernandez (2006) demonstrated how powerful choice can be in a study with six pre-school children. Their results indicated that children preferred having choice in their reinforcer, even when reinforcers were held consistent in the choice and

no choice conditions. They also reported that students continued to choose the task that allowed them choice in their reinforcement even when the choice-based tasks required more work than the no-choice tasks. According to a meta-analysis by Shogren and colleagues (2004) choice has consistently demonstrated positive effects on student behavior. Although there is a dearth of research examining choice-based interventions for teachers, there is some research in other fields to suggest that it may also be an effective antecedent method for improving adult behavior. Javenic and colleagues (2003) conducted a study to determine if participants' choice in the instructional format (self-directed vs. group) in which they received health education would result in a change in adult behavior. Results indicated that individuals who had the opportunity to choose the instructional format were more likely to attend sessions compared to those who were randomly assigned to a teaching format. Based on the strong literature base for choice-based interventions, there is evidence to suggest that choice may serve as a powerful reinforcer for individuals. Thus, a teacher's preference for an intervention to implement may serve as a powerful reinforcer, resulting in increases in treatment integrity.

Purpose

Problem-solving consultation is an effective method frequently used by school psychologists to identify strategies and train teachers to implement evidence-based interventions for academic, behavioral, and mental health concerns (Sheridan, Welch, Orme, 1996; Gutkin 1996; Kratochwill & Stoiber, 2000). The problem-solving consultation framework is missing a key piece for implementing effective interventions in the school, the evaluation of implementation procedures. TI is defined as the degree to which an intervention is implemented as it is intended (Gresham, 1989). When evaluating

intervention effectiveness, it is necessary to evaluate both student outcomes and TI. Evaluations made without TI data may be invalid and lead to inaccurate treatment decisions for students. The literature suggests that TI is often not reported in the literature nor is it often assessed in practice (Sanetti et al., 2011). This is concerning, as research evaluating TI has consistently demonstrated that teachers often implement interventions with poor integrity (Noell et al., 2005). Thus, it is necessary to identify methodologies that are effective for improving the integrity with which interventions are implemented.

Strategies to promote TI are primarily responsive methods that allow teachers to implement interventions with low integrity before providing supports. Examples of these methods include performance feedback, self-monitoring, motivational interviewing, participant modeling, role playing, and implementation planning. Although these methods have been demonstrated to be effective for improving teacher TI, high levels of integrity are often not maintained (Solomon, Kelin & Politylo, 2012). Most importantly, the majority of these strategies have been implemented as responsive methods to poor TI. Additional strategies are needed to serve as preventative strategies, decreasing the likelihood that low TI ever occurs. One factor that should be considered for improving TI is treatment acceptability (Allinder & Oats, 1997). Teachers may be more likely to utilize an intervention if they find the intervention acceptable and its performance of the intervention matches their expectations. The test-driving intervention that incorporates teacher acceptability and choice may be one option for a preventative method that has been empirically evaluated (Dart et al., 2012).

Test-driving is an intervention for teachers to 'test' or briefly implement multiple strategies to address student referral concerns. Teachers are then given a choice to select

their most preferred intervention, which is assumed to be the strategy that best matches their expected performance and implement this intervention consistently. Although the test-driving intervention was effective for promoting TI across teachers, no data were collected regarding the procedural integrity of the intervention, nor for student outcome data. Additionally, the design of this study did not meet WWC standards, limiting the external and internal validity. Thus, it is unknown whether test-driving was responsible for the improvements in integrity data. The purpose of the current study is to further evaluate the effectiveness of test-driving for teachers by collecting data during the test drive conditions to determine whether the "test drive" was responsible for the improvements observed in the level of TI.

Research Questions

1. Will teachers who are non-adherent to interventions following traditional problem-solving consultation (i.e., demonstrate levels of TI of 50% or lower) demonstrate changes in TI during the teacher-choice intervention phase following a test drive phase?

2. Will there be a relationship between TI and student academically engaged behavior?

3. Will there be a relationship between teacher's ratings on the URP-IR for the acceptability, understanding, and feasibility factors and teachers' most preferred intervention?

4. How will teachers perceive the test drive intervention on the URP-A?

CHAPTER II - METHOD

Participants and Setting

The study was conducted at a rural high school located in the Southeastern United States. There were 588 students enrolled in the school, with approximately 60% of students qualifying for free and reduced lunch. According to the school's data, 68% of students identified as White, 29% identified as Black, 2% identified as Hispanic, and 1% as two or more races. Prior to the study, approval was obtained from the university's Institutional Review Board (IRB), school administration, participating teachers, and student's parents/guardians. Additionally, informed consent was obtained teachers, students, and parents/guardians prior to implementation of the intervention. Three teacher-student dyads were recruited based on teacher referral to the school's behavioral consultants for concerns regarding student behavior. The teachers were interviewed by the school's behavioral consultant to identify the presenting problem behaviors and to obtain basic demographic information. Consent was obtained from all teachers, parents/guardians, and students prior to implementation of the intervention due to the collection of individual student and teacher data; however, teachers were only included in the study if they demonstrated TI levels of less than 50% across three consecutive screening sessions. The first three teachers who referred students to the behavioral consultants met this criterion and were thus recruited for participation in the study.

Teacher/Student Dyad One

Stephen identified as a 15-year-old White male student and was enrolled in the 10th grade. His English teacher, Caroline, submitted a referral to the behavioral consultants for talking out of turn in class. Stephen was absent a total of 19 school days

(16 of which were due to mandatory quarantines following direct exposure to COVID-19) and he did not have any office discipline referrals (ODRs). During the brief student interview to determine his preferences for reinforcement, he requested to earn headphones. Since these were a larger more expensive item that could not be provided following implementation of intervention each day, he opted to earn tokens that he could exchange for $5 Walmart gift cards to put towards his headphone purchase. He needed three tokens to exchange for a gift card and could earn one token per day of intervention. Stephen's teacher, Caroline, was a 31-year-old White female with eight years of teaching experience. She held a master's degree in education and primarily taught English courses. At the time of the study, there were a total of 15 students enrolled in the class, with the majority of students identifying as male (80%), and one student who qualified for special education services under the Other Health Impairment (OHI) category.

Teacher/Student Dyad Two.

Damon identified as a 15-year-old White male and was enrolled in the 9th grade. He was referred by his English teacher, Lexi for concerns related to off-task behavior. A brief review of Damon's records indicated that he had four ODRs for the following behaviors: profanity, defiance, inappropriate engagement with peers, and being out of area. The student was absent for a total of 16 school days during the study (10 of which were due to a required quarantine following direct exposure to coronavirus). Based on a brief interview with the student, Damon indicated that he would like to work for preferred snacks (i.e., chips or soda). Damon's teacher, Lexi was a 38-year-old White female with 10 years of teaching experience. She held a Bachelor of Arts degree and primarily taught English courses. There were a total of 23 students enrolled in her class, with 56% of students

identifying as male, and 30% of students qualifying for special education services under various categories including: Emotional Disability (EmD), Specific Learning Disability (SLD; including both reading and math) and OHI.

Teacher/Student Dyad Three.

Elena was enrolled in the 10th grade and identified as a 17-year-old White female. She qualified for special education services under the SLD category due to deficits in basic reading, reading comprehension, and reading fluency categories. She was referred by her inclusion teacher, Bonnie for concerns of off-task behavior in her English class. Elena did not have any ODRs and was absent a total of 33 school days (25 of which were due to mandatory quarantines following close contacts to individuals who tested positive for COVID-19 and due to the presence of COVID-19 symptoms). During the brief student interview, Elena identified sour candy as her preferred items for reinforcement. Elena's teacher, Bonnie identified as a 27-year-old White female with 3 years of teaching experience. She held a Bachelor of Art's degree and was the inclusion teacher for Elena's English class. The English class consisted of 17 students, with 76% of students identifying as male and 35% of students qualifying for special education services under various categories including: EmD, SLD (both reading and math), and Speech/Language impairment.

<div align="center">Materials and Measures</div>

Treatment Overviews and Integrity

Five brief overviews of each treatment option (i.e., self-monitoring, response cost, modified Check-in/Check-out (CICO), behavior specific praise (BSP), and pre-correction) were developed for teacher trainings (see Appendix B). In addition to this,

five integrity checklists were developed and included the necessary components for accurate implementation of each intervention (See Appendix A). A rationale and brief description for each treatment option is described in detail in the Interventions section below.

Rewards

Items for reinforcement were determined by a brief preference assessment conducted during a student interview. These items were provided to students if they met the reinforcement criteria at the conclusion of the intervention. Rewards were screened and approved by the students' classroom teachers and are listed above in the participant section.

MotivAider®

A MotivAider® is a small device that can be used to discreetly prompt an individual to engage in some response. The device can be set to provide a tactile prompt (i.e., vibration) on a fixed or variable schedule. For the purposes of this study, the MotivAider® was utilized for the self-monitoring intervention. The MotivAider® provided a prompt on a fixed-interval schedule to prompt a student to engage in self-monitoring behaviors or to prompt a teacher to provide BSP.

Usage Rating Profile – Intervention, Revised (URP-IR)

The URP-IR is a 29-question rating scale that can be completed by teachers to determine the extent to which they found the intervention as socially valid. Individuals rate their agreement with each item on the scale from one (strongly disagree) to six (strongly agree). The scale assesses six different factors of the social validity construct, including Understanding, Acceptability, Family-School Collaboration, Feasibility,

System Climate, and System Support (Briesch, Chafouleas, Nugebauer, & Riley-Tillman, 2013). For the purposes of the present study, the rating scale was be adapted to only include items related to Acceptability and Feasibility. These factors most closely aligned with the purpose of the intervention and only included 17 items, which was more feasible for teachers to complete following the implementation of each intervention. Alpha coefficients for the selected factors were reported as .96, .80, and .84. Higher scores on each of these factors will indicate greater acceptability and/or satisfaction of the intervention.

Usage Rating Profile – Assessment (URP-A)

The URP-A is a 28-question rating scale that can be completed by teachers to evaluate the extent to which they found an assessment procedure acceptable, feasible, and useful (Chafouleas, Miller, Briesch, Neugebauer, & Riley-Tillman (2012). The URP-A was used to determine whether teachers found the test-driving intervention procedure acceptable. The scale assesses the same six factors listed in the URP-IR rating scale (i.e., Understanding, Acceptability, Family-School Collaboration, Feasibility, System Climate, and System Support). Raters mark their agreement on a scale of one (strongly disagree) to six (strongly agree) and results are calculated by adding the ratings for each factor and dividing by the total number of questions in each factor. Higher scores on each of the factors indicates more acceptability, feasibility and usefulness of the assessment tool, with the exception of the System Support factor. A higher score on the System Support factor indicates that the teacher would need a greater level of support in order to implement the assessment independently. Alpha coefficients for the factors on the URP-

A were between .63-.90 (Miller et al., 2014). No items or factors were be excluded from the URP-A.

<center>Dependent Measures</center>

The primary dependent variable was TI, or the extent to which teachers implemented the behavioral interventions with fidelity. TI was assessed based on teachers' adherence to the treatment components detailed in the aforementioned integrity checklists for each of the five interventions included in this study. TI was reported as the number of intervention steps implemented divided by the total number of intervention steps possible and multiplied by 100 to report as a percentage. Secondary dependent variables included students' academic engaged behavior (AEB), disruptive behavior (DB), and passive-off task behavior (POT). AEB was defined consistent with previous studies (e.g., Briesch, Chafouleas, & Riley-Tillman, 2010). A student was coded as AEB if they were actively attending (e.g., asking a question, participating in class discussion, or writing) or passively attending (e.g., reading, oriented towards the task or teacher) to the current academic task demand. Definitions for DB were defined as behaviors that distracted or interfered with the students' ability to attend to the task demand. DB was further defined for each student based on teacher reports. Identical definitions were developed for Stephen and Damon (dyads one and two) and included talking out of turn, inappropriate vocalizations, and out of seat behavior. Examples of DB for these students included talking about non-academic related tasks, blurting out answers without permission, and walking around the classroom without permission. DB for Elena included all components of the above definition and inappropriate use of technology. Examples included playing games on her Chromebook, use of cellphone or headphones,

<center>29</center>

and opened websites unrelated to the current task demand (e.g., facebook).POT was defined consistent with the Behavioral Observation of Students in Schools (BOSS; Shapiro, 2004) coding scheme for off-task passive and included behaviors such as sleeping and looking away from the current task demand.

Data Collection

Data regarding adherence to TI and student behavior were collected during 15-minute sessions using systematic direct observation (SDO). TI was assessed by observers using the TI checklist that coincided with the intervention being implemented (see Appendix A). Student behavior (i.e., AEB, DB, and POT) was assessed using 10-second momentary time sampling. For each observation session, observer(s) entered the room and stood in an unobtrusive location in the back of the classroom for a minimum of five minutes prior to the start of the session. At the beginning of each 10s interval, the observer looked at the target student, determine whether they were engaged in AEB, DB, or POT and recorded the respective behavior on the data sheet. At the end of the 15-minute session, the observer calculated the percentage of intervals in which the student was engaging in AEB, DB, and POT by dividing the number of intervals they were engaged in each by the total possible number of intervals (i.e., 90) and multiplying by 100. TI was also collected throughout the observation period, as components of the intervention were implemented.

Observer Training and Interobserver Agreement (IOA)

Seven observers, five graduate students in a doctoral school psychology program and two undergraduates recruited from a university school-psychology based lab, were trained on systematic direct observations using behavioral skills training (i.e.,

instructions, modeling, rehearsal and feedback). Observers watched a 10-minute video in which they were required to meet 90% agreement using an interval by interval method of calculation for student behaviors including academically engaged, disruptive, and passive off task behaviors. All observers met this criterion on their first attempt. A brief didactic training was then provided regarding the independent and dependent variables, operational definitions, and data collection procedures (e.g., 10-second momentary time sampling) of the present study. Total training time lasted approximately 45 minutes.

A secondary observer was present during an average of 47.67% of observation sessions (range 33.33-60%) for each teacher-student dyad across the consultation-as-usual and teacher-preferred phases for both TI and student behavior. IOA for procedural integrity was also assessed for 50% of teacher/student dyad two's test drive phase but was not assessed for teacher/student dyad one due to resource constraints (i.e., time restrictions and lack of available observers). Exact IOA (Cooper, Heron, & Heward, 2020) was calculated using a component-by-component method for TI and interval by interval method for student behavior. To calculate IOA using this method, the total number of components or intervals that the observers agreed on were divided by the total number of components or intervals possible. Secondary observers fell below 90% agreement on one occasion for student behavior and a brief retraining (i.e., corrective feedback and a review of procedures) was completed before future observations were conducted. IOA for treatment integrity was 100% for teacher/student dyad one across all phases, 97% (range 75%-100%) for teacher student dyad two across all phases, and 100% for teacher student dyad three for the consultation-as-usual phase. IOA for student behavior averaged at 97% (range: 88.52%-100%) for teacher/student dyad one, 94.44%

(range: 90.74%-96.67%) for teacher/student dyad two and 98.15% (range 96.30%-99.63%) for teacher/student dyad three.

<div align="center">Interventions</div>

The interventions that were selected for the purpose of this study included self-monitoring, modified Check-in/Check-out, response cost, behavior specific praise, and precorrection. These interventions were selected based on their strong evidence-base, theoretical consistency, similar number of components necessary for implementation, and similar resources required for implementation. The interventions were slightly adapted to fit the 15-minute data collection session to ensure that students had the opportunity to earn reinforcement during the session. Refer to Appendix B for an overview of each intervention and its associated materials.

Self-Monitoring

Self-monitoring has been an effective strategy for improving on-task behaviors in the classroom setting (Amato-Zech, Hoff, & Doepke, 2006). This intervention required a self-monitoring form and a MotivAider®. The teacher provided the MotivAider® to the student, which was set to three-minute fixed intervals, and provided brief instructions to the student about how to complete the self-monitoring form. Each time the MotivAider® vibrated, the student placed a check mark or an "X" on the self-monitoring form to indicate whether they were on or off task at the time of the vibration. The teacher also closely monitored the student's behavior so that they could conduct an agreement check at the end of the observation. If the student's ratings matched the teacher's ratings with 80% accuracy, and the student indicated that he or she was on-task for at least 80% of the

observed intervals (i.e., 4 of 5), then they were be provided access to their previously identified reward.

Modified Check-in/Check-out (CICO)

CICO has been utilized in the schools to address many different behaviors in the school setting (Hawken & Horner, 2003; Todd, Campbell, Meyer, & Horner, 2008). It is traditionally used to check-in with students when they first arrive to school, and check-out with students at the end of the school day. For the purposes of the study, the CICO procedures were modified to fit the 15-minute observation period. Teachers checked-in with the student at the beginning of the observation period by providing encouragement, three behavior expectations, and a behavior goal required to access reinforcement. At the end of the 15 minutes, the teacher completed a direct behavior rating (DBR; Gresham, 2010; Kilgus, Riley-Tillman, Chafouleas, Christ & Welsh, 2014) of the student's on-task behavior. Teachers then checked-out with the student by providing them corrective feedback and encouragement. If students earned 80% or more of their points on the DBR, they earned access to their previously identified reward.

Response Cost

Response cost is a strategy that has been used to decrease disruptive behaviors in the school (Tiano, Fortson, McNeil, & Humphreys, 2005). For the response cost procedure, a stack of five sticky notes were placed on the student's desk. Every time the student engaged in a disruptive behavior, the teacher removed a sticky note from the pile. If 80% or more of the sticky notes remained at the end of the 15-minute session (i.e., four out of five sticky notes) the student was provided with their previously identified reward.

Behavior Specific Praise (BSP)

BSP can be defined as a statement teachers provide to students that pairs praise (e.g., "good job") with a specific behavior (e.g., "sitting in your seat,"). BSP is typically used as a form of positive reinforcement. When teachers observe students engaging in appropriate/desired behaviors, they may use BSP in order to increase the future likelihood that the behavior continues to occur. BSP can be provided to a group for their collective behavior or to an individual. BSP has been utilized to promote positive behaviors in the classroom for students of all ages (Haydon et al., 2020). Although there are differing suggestions on the number of praise statements to provide to students, literature has suggested providing one praise statement per two minutes. Due to already low levels of praise statements, teachers in the current study were recommended to provide four BSP statements during the 15-minute session (approximately one statement every three minutes) and were recommended to ignore minor disruptive behaviors to maintain a 4:1 ratio.

Precorrections

Precorrections or prompts are utilized to remind students of behavioral expectations. Precorrections are often used prior to task demands that have previously incited behavioral difficulties. For example, if students typically have trouble transitioning in the hallway, teachers may remind students of the expectations for the hallway (e.g., "remember to stay on the right side of the hallway, use a quiet voice and keep your hands and feet to yourself"). Precorrections have been successful in reducing problem behaviors in a variety of settings, especially when paired with active supervision (De Prey & Sugai, 2002). At the beginning of the session, teachers briefly met with the

student to review their behavioral expectations for the next 15 minutes. Halfway through the observation, they briefly checked in with the student to remind them of the rules. Teachers also provided praise when the students were engaging in the expected behaviors and re-direction when students were not engaging in the expected behaviors.

Design

A single-case multiple baseline design across three teacher-student dyads was utilized to evaluate the effects of the test-driving intervention. Teacher-student dyads were yoked to help further eliminate any threats to the internal validity of the study by acting as control participants. Teacher-student dyad two was yoked to teacher student dyad one, and teacher student dyad three was yoked to teacher-student dyad two. WWC standards for multiple baseline design include (a) a minimum of five data points per phase, (b) a minimum of 80% IOA for at least 20% of observations across participants and phases, (c) a minimum of three replications, and (d) systematic manipulation of the independent variable. Data were staggered by at least two sessions across all participants. Because of the nature of the test drive procedure, WWC standards were not met during the test drive phase. The test drive phase only contained four independent data points, which represented implementation of the remaining interventions that were not implemented during the consultation-as-usual phase. Teacher/Student dyad three was unable to complete the test-drive procedure due to multiple mandatory quarantines from the COVID-19 pandemic. Thus, WWC standards were not met as three replications were not observed. All other WWC standards listed above were met.

Data Analysis

Data were primarily analyzed via visual analysis of level, trend, and variability. Additional considerations for treatment effects included analyzing the nonoverlap, immediacy, and consistency of data. To minimize the potential for misinterpretation during visual analysis of data, graphs were constructed with single-case graph guidelines in mind. The ordinate axis was scaled from 0-100%, which encompassed all possible ranges of percentages for student behavior and TI and the data points per x- to y-axis ratio (DPPXYR) was calculated at .25, which is within recommended guidelines of .14 or larger (Radley, Dart & Wright, 2018). All phase change decisions were made based on the primary dependent variable, TI. Single-case effect sizes, Baseline Corrected Tau (BCT; Tarlow, 2017) was calculated. BCT was chosen over other methods, such as Tau-U because it accounts for trends in addition to non-overlap of data. BCT also utilizes the median in its analyses rather than the mean, which makes BCT a more conservative measure of effect size, as it is not impacted as strongly by outliers in the data. If there were no statistically significant trends in baseline, data were reported as Tau values, which are still bound between -1 and +1.

Procedures

Consultation-As-Usual

The primary researcher provided a brief didactic training to two of the school's behavioral consultants (i.e., graduate students enrolled in a school psychology doctoral program) on the procedures of the study. The two school-based consultants assumed the role of the behavioral consultant and followed the problem-solving consultation model for student referrals. Once teachers made a referral to the behavioral consultants, the

teacher and behavioral consultant met and followed the steps outlined in the problem identification stage. During this stage, they clearly identified a target behavior and selected a 15-minute block to implement the intervention, which was the time the teacher reported as most problematic. Next, the behavioral consultant and teacher began the problem analysis stage where information from problem identification stage was utilized to select the most appropriate intervention from the list above (i.e., BSP, modified CICO, pre-corrections, response cost, and self-monitoring). For example, teacher one reported that the student typically responded well to prompts to stay on-task, but that it was difficult to continuously provide prompts while she was working with other students. Thus, a self-monitoring intervention closely matched the needs of both the teacher and the student. Teacher two stated that she felt the student would benefit from more tangible reinforcers in comparison to social reinforcers such as praise. Based on these reports, consultants narrowed the intervention selection down to response cost and modified CICO. Because there were no clear indicators for which intervention would be more appropriate, they randomly selected response cost. Teacher three indicated that she had a great relationship with her student and wanted to find ways to help mentor her. Based on this information, the modified CICO intervention seemed to be the best fit.

Next, teachers were trained on the recommended strategy using behavioral skills training. Training lasted approximately 10-15 minutes and was conducted in-person during their designated planning period. The consultants first provided a brief introduction and rationale for the intervention, modeled the intervention, and then allowed the teacher the opportunity to rehearse the intervention. The consultants then provided corrective feedback and verbal praise until the teacher demonstrated the

intervention with 100% accuracy within the training setting. Following accurate demonstration, the teacher began the plan implementation phase of problem-solving consultation, where they were responsible for implementing this strategy for a minimum of five sessions without feedback form the consultants. The behavioral consultants in conjunction with other trained observers (both undergraduates and graduate students) collected data on both TI and student outcomes during each session during the plan implementation stage. Teachers who demonstrated an average of 50% or lower levels of TI across all sessions, were recommended to continue with participation in the study. Teachers who demonstrated an average of 51% or more TI would not have been recommended for participation and would have been provided with typical behavioral consultation, including other evidence-based procedures consistent with current practices (e.g., performance feedback) All recruited teachers demonstrated 50% or less integrity across the sessions and were therefore recommended to participate.

Test-Drive

The first teacher-student dyad began the test drive phase following a minimum of five data points in the consultation-as-usual condition. Teacher-student dyads two and three remained in the consultation-as-usual phase while teacher-student dyad one completed the test drive intervention. Teachers were informed that they would be "test-driving" four different interventions that have been demonstrated as effective for reducing problem behaviors. The remaining interventions that were not implemented during the consultation-as-usual or the yoked-intervention phase were written on slips of paper and placed in a bag. The behavioral consultants determined the order of the interventions for the test drive by randomly selecting a piece of paper out of the bag.

38

Teachers were trained on two of the four intervention procedures at a time and were responsible for implementing both interventions in one class period. After implementation of the first two interventions, procedures were replicated for the second two interventions. Decisions to implement two interventions per day in this phase were made based on time constraints due to the COVID-19 pandemic and required state testing. Training for intervention procedures were held consistent with the consultation-as-usual phase (i.e., BST). Following training, teachers were required to implement each intervention in the test drive with 100% adherence to intervention procedures. Treatment integrity collected during this phase also served as procedural integrity, which was collected for 100% of test drive sessions. The checklist was utilized to determine whether all interventions were implemented with 100% integrity during the test drive phase. Procedural integrity was calculated by dividing the number of steps implemented by the total number of possible steps and multiplied by 100. To increase the likelihood that teachers implemented the interventions with 100% integrity during this phase, teachers were offered a small gift for completing the checklist. A brief preference assessment via teacher interview was utilized to determine each teacher's reward (i.e., $25 gift card). Teachers accessed the reward if 100% TI was observed on the first implementation of each intervention. After test-driving all interventions, the teachers completed items on the URP-IR related to acceptability, feasibility, and understanding for each intervention in the test drive phase. Additionally, teachers separately rank-ordered the interventions in the test drive phase that were implemented with 100% integrity from their most to least preferred intervention. It is important to note that due to time constraints, the teacher in dyad two was not required to re-implement the fourth intervention despite TI below

39

100%. Due to this, the intervention was removed from potential choices during the teacher-preferred stage.

Preferred Intervention

After the test-driving phase, teachers implemented the intervention that they ranked as their most preferred intervention. A minimum of five data points were collected during this phase for all teachers. No feedback was provided following any sessions during the preferred intervention phase.

Yoked Intervention

During the yoked-intervention phase, it was planned that the teacher in dyad two would implement dyad one's preferred intervention and the teacher in dyad three would implement dyad two's preferred intervention. The yoked intervention phase was added to strengthen internal validity and ensure that the test-drive procedure (and not the specific intervention) was responsible for any observed changes in TI. Because teacher/student dyad two implemented teacher/student dyad one's most preferred intervention during their consultation-as-usual phase, a yoked intervention phase was not necessary for teacher/student dyad two. If time allowed, teacher-student dyad three would have begun the yoked intervention phase in which they implemented teacher/student dyad two's most preferred intervention. Following a minimum of five data points during this phase, and a continuation of 50% average or lower TI, teacher-student dyad three would have completed the test drive phase, as described above, with the remaining four interventions, and moved to the teacher-preferred phase.

Social Validity

At the conclusion of the study, teachers were asked to complete the URP-A to determine the extent to which they found the test drive intervention acceptable.

CHAPTER III - RESULTS

Teacher/Student Dyad One

The results of teacher/student dyad one can be found in the top panel of Figure 1 in Appendix G. During the consultation-as-usual phase, TI data were at low to moderate levels, with little variability (average 40%, range: 25-50%). Data for student AEB were at high levels with an overall decreasing trend (average 78%, range: 62.22 - 95.56%). During the test drive phase, the teacher implemented three of the four interventions with 100% integrity and the fourth intervention with 40% integrity, During the test drive phase, there was an immediate increase in stability for student AEB and levels remained very high (average 88%, range: 82.22%-90%). Once the teacher-preferred phase was implemented, TI data decreased to moderate to low levels with some variability (average 52.78%, range: 33.33-75%), which were comparable to baseline levels. Although levels of student AEB remained high during this phase, an overall slight decreasing trend was again observed across time (average 87.75, range: 68.89-98.89%).

As a secondary analysis, baseline corrected tau values were calculated to compare levels of TI and AEB from consultation-as-usual to teacher preferred phases. When levels of TI were compared from consultation-as-usual to the teacher preferred phase, a small effect was observed (*Tau* = .320). A small effect was also observed for student AEB (*Tau* = .323). These data align closely with reports from visual analysis data.

Teacher/Student Dyad Two

The second panel of Figure 1 represents data from teacher/student dyad two. During the consultation-as-usual phase, TI data started at moderate levels of implementation, and after six sessions a clear decreasing trend in TI was observed

(average 50%, range: 25-75%). Student AEB during this phase was at low to moderate levels with an increasing trend prior to implementation of the intervention (average 40.56%, range: 6.67-65.56). There were no observed differences to TI or student AEB during the consult-as-usual phase following teacher/student dyad one's implementation of the test drive phase nor during the teacher-preferred stage. During the test drive phase, the teacher in dyad two implemented all four randomly selected interventions with 100% accuracy. No immediate changes were observed in student AEB during the test drive phase; however overall level was higher (average 68.22%, range: 48.89-81.11%). During the teacher-preferred phase, TI data was observed at higher levels compared to the consultation-as-usual phase, and data were more variable (average 65%, range: 25-100%). Student AEB also remained variable and was observed at an overall higher level when compared to consultation-as-usual (average 66%, range: 36.67-84.44%).

BCT values were also calculated for dyad two and were compared from the consultation-as-usual phase to the teacher-preferred phase for both TI and student AEB data. Results indicated that there were small effects observed in TI data (Tau = .201) and student AEB data (Tau = .444), which is consistent with reports of visual analysis.

Teacher/Student Dyad Three

Results of teacher/student dyad three are displayed on the third panel of figure 1. During the consultation-as-usual phase, TI data were variable and at low to moderate levels with a slight increasing trend in the final few data points (average 34%, range 0-75%). Student data for AEB were high with little to no variability (average 92.67%, range 85.56-98.89%). No effects were observed during the consultation-as-usual phase

during teacher one's implementation of the test drive and preferred intervention phases. Additional data was unable to be collected and analyzed due to resource constraints.

<center>Acceptability and Social Validity Data</center>

Acceptability of Interventions

All teachers completed an adapted version of the URP-IR which only included items related to factors of acceptability and feasibility. Teachers completed the adapted URP-IR at the end of the consultation-as-usual phase and after each intervention during the test drive phase. Data was compiled from all teachers and was reported by teacher, intervention, and factor. These results can be viewed in Appendix G, Table 1. Caroline, the teacher in dyad one, rated all interventions on the URP-IR with high acceptability and feasibility. During the Teacher's Most Preferred phase, Caroline selected Self-Monitoring as her top intervention choice, which aligns with her top ratings on the URP-IR. Lexi, the teacher in dyad two, had more variability in her ratings in comparison to Caroline. According to her responses, she perceived Response Cost as the most acceptable intervention and CICO and BSP as the most feasible interventions. Caroline's ratings for both acceptability and feasibility for her Most Preferred intervention (response cost) were high in comparison to her ratings for other interventions.

Social Validity of Test-Drive Procedures

Teachers one and two also completed the URP-A at the end of the study regarding their perceptions of the test-driving procedures. Based on results of the URP-A, teachers' perceived the test-driving intervention as acceptable. Average ratings for each factor can be found in Table 2 of Appendix G. Results of each factor indicated that teachers found

<center>44</center>

the test-driving procedure as acceptable, understandable, feasible, and fitting with their system climate. Some differentiation was identified for how each teacher rated items related to system support and home/school collaboration. Teacher one's ratings indicated only some home/school collaboration would be needed, while teacher two indicated greater home/school support would be necessary. Additionally, teacher one indicated they would need additional support from school administration or other staff in order to implement the test-driving procedures, while teacher two indicated needing little to no support to implement these procedures.

CHAPTER IV – DISCUSSION

Problem-solving consultation is an effective method frequently used by school psychologists to identify strategies and train teachers to implement evidence-based interventions for academic, behavioral, and mental health concerns (Gutkin, 1996; Kratochwill & Stoiber, 2000). During the treatment evaluation stage of problem-solving consultation, evaluations of the recommended interventions should include both student response to treatment and treatment integrity (TI) data (Shadish, Cook, & Campubell, 2002). TI is directly related to student outcomes, meaning greater levels of TI typically result in greater levels of student improvement (Noell, Gresham, & Gansle, 2002). In the majority of research, it is suggested that TI levels are either a) not reported, b) are reported with very low levels, or c) are initially reported high but decrease within a few days of implementation (Sanetti, Dobey, & Gallucci, 2014).

Strategies to promote TI are primarily responsive methods, which may be considered "wait to fail" models. Preventative methods are needed to decrease the likelihood that low TI ever occurs. To date, only one preventative intervention (i.e., test-driving) has been empirically evaluated (Dart et al., 2012). Test-driving is an intervention for teachers to 'test' or briefly implement different strategies to address student referral concerns. After each intervention has been tested, teachers select their most preferred intervention and implement it consistently across time. Test-driving may be an effective method for improving TI, as it incorporates aspects of treatment acceptability and choice. Although the test-driving intervention was effective for promoting TI across teachers, there were many limitations of the study. Thus, it is unknown whether test-driving was responsible for the improvements in integrity data. The purpose of the current study was

to further evaluate the effectiveness of test-driving using a single-case multiple baseline design across three teacher-student dyads. Results surrounding each research hypothesis are described below.

<div align="center">Research Question One</div>

The first research question aimed to evaluate changes in TI data from the consultation-as-usual-phase to the teacher-preferred phase. Based on the results of both visual analysis and the secondary effect size calculations, minimal changes were observed in TI data. Although a slight increase in the average level of TI was observed for the teachers in the first and second dyad, differences may not be considered clinically meaningful. These results are inconsistent with Dart and colleagues (2012) initial evaluation of the test drive intervention, where immediate increases were observed in TI data for all four participants. Results may differ from the previous study for various reasons. There were very variable levels of TI observed for all participants during the current study, in comparison to lower more stable levels in the previous study. Additionally, there were fewer participants included in the study, therefore it is difficult to determine potential external validity. The current study was also conducted in a high school setting, where research on interventions is already limited in number in comparison to elementary and middle school participants, which were utilized in the previous study.

There are some other potential explanations for moderate levels of integrity following the test drive phase. First, both teachers selected interventions with tangible reinforcers to implement in the teacher-preferred stage; however, both teachers reported that they delayed the delivery of reinforcement to the end of the class period due to

concerns of student problem behavior. Because reinforcement was not delivered within

five minutes of concluding the intervention, and therefore was not observed by data

collectors, this component was consistently marked as incorrect on the data sheet.

Second, both teachers often missed the step of reminding students of the components of

the intervention prior to beginning implementation. It is possible that teachers felt

additional explanations or reminders of key components of the intervention were

unnecessary for high school students. These data are consistent with previous research,

which has indicated that although teachers may initially implement interventions with

high integrity (e.g., the test-drive phase), TI typically drops between levels of 0-65%

within the first ten days of implementation (e.g., preferred intervention phases;

Mortenson & Witt, 1998; Noell, Witt, Gilbertson, Rainer, & Freeland, 1997; Sanetti &

Kratochwill, 2008).

<p style="text-align:center">Research Question Two</p>

Another purpose of the study was to investigate the relationship between TI and

student AEB. No relationship was identified between the two variables. In the first phase

for teacher/student dyad one, student AEB did fluctuate with percent of adherence to

treatment components. Additionally, when the test drive was implemented and 100% of

treatment components were implemented, student AEB was high and stable. Based on

these data, it was expected that student AEB would have dropped to lower levels during

the fourth intervention that was implemented during the test-drive due to low levels of TI,

however it remained high and stable. There was no clear relationship between TI and

student AEB during the teacher-preferred phase for dyad one. There also was no clear

relationship between these variables for any phases within dyad two. Even when

<p style="text-align:center">48</p>

interventions were implemented with 100% integrity during the test drive phase for student two, student behavior remained variable. Due to the variability of student AEB and TI during the consult-as-usual and teacher-preferred stages, no functional relationship was identified.

These results are also inconsistent with previous research. In the original test drive study, a strong positive relationship between student AEB and TI was observed for three of the four participants (Dart et al., 2012). Other research investigating relationships between TI for academic interventions and student performance has demonstrated that increased levels of TI are associated with improved student outcomes (Noell, Gresham, and Gansle, 2002). A relationship may have been difficult to derive due to already high levels of AEB for dyad one. There are also other potential factors that may have a greater influence on student behavior than adherence to treatment components. A student's behavior may change based on how the intervention is presented (e.g., tone of voice, facial expressions), student/teacher relationship, potency of reinforcers, function of behavior, and other environmental factors. These factors may even be exacerbated at the high school level due to more complex skills and behaviors as well as a difference in classroom environments when compared to the elementary and middle school level.

<div align="center">Research Question Three</div>

Third, the study aimed to determine whether there was a relationship between ratings on the adapted URP-IR and the selection of the teacher-preferred intervention. Both dyads that completed the test drive selected interventions in the teacher-preferred phase that they also rated high in acceptability and feasibility on the URP-IR. There were very clear differences in Caroline's ratings on the URP-IR across interventions, with her

preferred intervention being consistently rated high for both the acceptability and feasibility factors. Lexi's ratings were more difficult to detect differences, as there were many interventions with high ratings on both factors. Two interventions were tied for acceptability and feasibility scores: self-monitoring and BSP. This is important to note because BSP was the intervention that was implemented with low integrity during the test drive, and thus was excluded from the list of potential interventions to implement during the teacher-preferred phase. If the teacher were provided with the opportunity to re-implement the BSP intervention, it is possible that her ratings on the adapted URP-IR and her selection for the teacher-preferred phase may have changed. Overall, ratings on the adapted URP-IR and teacher's choice of their most preferred intervention were aligned. Despite alignment of the adapted URP-IR ratings and the selected most preferred intervention, the data suggests that acceptability and/or choice was not sufficient to improve TI data. Some research has previously identified a relationship between acceptability and TI (Allinder & Oats, 1997) although other research has reported these variables as unrelated (Noell et al., 2005). Additional research is needed to continue to evaluate the relationship between TI and acceptability/choice.

Research Question Four

Finally, the study aimed to evaluate how teachers perceived the test drive intervention as indicated on the URP-A. According to teachers results on the URP-A, both teachers who implemented the test-drive intervention rated it as an acceptable procedure, with consistently high scores across factors of acceptability, understandability, feasibility, and fit with school climate. Differences were observed in the ratings for home/school collaboration and system support, with teacher one indicating a need for

greater support from school staff when compared to teacher two. It may be unsurprising that a teacher may need greater support to implement the test-driving intervention, as it requires implementation of multiple different interventions across a few days. Additionally, problem-solving consultation itself utilizes school resources (e.g., behavioral consultants, school psychologists) that teachers may need access to implement the test-driving intervention. Teachers may need assistance from these school-based personnel to select appropriate interventions and access resources required to implement the interventions (e.g., MotivAider®, rewards). The second difference was observed on the home/school collaboration factor, with teacher two indicating a greater need for home/school collaboration than teacher one. Perceptions of needing higher home/school collaboration may be due to a variety of factors. First, parent contact was required to obtain informed consent as well as to select appropriate student rewards. Second, teacher two reported consistent collaboration with the parent due to concerns with student behavior; thus, implementation of these interventions may have constituted further parent/teacher interactions to discuss student progress. Despite these differences in ratings on the URP-A, teachers generally viewed the test-driving procedure as acceptable. Acceptability data is important to determine whether stakeholders (e.g., parents, teachers, students), viewed an intervention or procedure as socially meaningful (Kazdin, 1977). In a recent review conducted by Silva and colleagues (2019), it was reported that almost two-thirds of intervention studies published from 2005 to 2017 did not include treatment acceptability data. Due to the mixed results examining the relationship between treatment acceptability and treatment integrity in the literature, it is important to include data

51

regarding acceptability and social validity in future studies to further evaluate this relationship.

Limitations

Readers should carefully consider the following limitations when interpreting results of the current study. The study was initially attempted in Spring 2020 and was halted due to the COVID-19 pandemic. Although schools returned in Fall of 2020, many restrictions were in place to promote health and safety in the school setting. Outside visitors and classroom observers were restricted and only permitted access to the school on a case-by-case basis. Restrictions were loosened mid-spring of 2021following the second wave of the pandemic, which is when the study took place. Although there were fewer restrictions in place, individuals who were identified as close contacts to those who tested positive with the COVID-19 virus and those who reported symptoms of the virus were required to quarantine for a minimum of 14 days. Thus, attendance of students, teachers, and data collectors was one of the largest impediments to the completion of the study that contributed to the many other limitations described below.

First, not all WWC standards were met. Due to fewer than three replications (i.e., only two completed participants), we were unable to draw accurate conclusions about the functional relationship between test-driving intervention and TI data. There are a few published studies that have drawn conclusions about a functional relationship with only two participants (e.g., Lindberg, Iwata, Roscoe, Worsdell, & Hanley, 2003); however, these data strongly supported their claims, unlike the high variability reported in the present study. Additionally, five data points were not collected per phase due to the nature of the test drive intervention, therefore this standard was met with reservations

(Kratochwill et al., 2010). Because all other phases met this criterion and more than three data points were collected during this phase, this limitation has minimal impacts on the internal validity of the present study. Standards related to IOA were also not met, as IOA was not obtained for the test drive phase for teacher/student dyad one due to limited resources. This limitation may also be considered minimal because a second observer was present for more than 30% of sessions and IOA was reported at high levels for both TI and student behavior, thus providing ample evidence for believability of the data.

Second, the first student referred to the behavioral consultants demonstrated fairly high levels of academic engagement, which potentially created a ceiling effect, making it difficult to determine any relationship between TI and AEB. Although this is a limitation, there are multiple studies that have evaluated the relationship between treatment integrity and student behavior and have found that higher levels of integrity leads to improved student behavior (Noell et al., 2002, Wilder et al., 2006).

Third, participants (all students and one teacher) were quarantined due to COVID-19 protocols at least once throughout the duration of the study. It is unknown how these large gaps in attendance may affect TI or student AEB. Fourth, Caroline in dyad one did not implement her fourth intervention during the test drive phase with 100% integrity. Thus, her pool of interventions to select from was smaller in comparison to dyad two. Fifth and finally, the current study did not incorporate student perceptions of the interventions. These perceptions would be important to consider as student input and buy-in are key components to the development of successful interventions.

Future Research

Future research should largely focus on the replication of the test-drive procedure, with emphasis on meeting WWC standards. It would be a disservice to analyze the results of this study without extreme caution, as there are valuable, evidence-based components to the test-driving procedure that warrant further evaluation. It is recommended that future researchers continue to investigate the relationship between choice, treatment acceptability, and treatment integrity and the relationship between treatment integrity and student outcomes. Rather than replicating the current study, future investigations may also look at a variety of ways to expand the current literature. One way might include obtaining student perceptions of the interventions and determining whether student preference matches teacher preference for interventions. Another suggestion may be to isolate the treatment acceptability and teacher choice components to determine if one method is more effective at promoting behavior change than combined methods.

Conclusion

TI and student outcome data are integral components to the evaluation of intervention methods. Despite emphasis on the importance of TI data, there are still a vast majority of studies that fail to report TI data, fail to collect TI data, or report low levels of TI data. Although there are a substantial number of evidence-based interventions for improving TI data (e.g., performance feedback, coaching) these methods have primarily been evaluated as responsive procedures. Preventative methods may be preferred to responsive methods, as they are proactive in nature. Test-driving may be a valuable preventative method for improving TI levels (Dart et al., 2012). Test-driving incorporates two components that may assist in the improvement of TI levels, treatment acceptability

and teacher choice. Test-driving is a procedure where teachers implement various interventions of similar effort and effectiveness to allow teachers the opportunity to align their expected outcomes of the intervention to reality before choosing a preferred intervention to implement across time. Initial evaluations of test-driving resulted in improved outcomes in TI data and student AEB. The current study aimed to expand upon the test-driving literature by replicating similar procedures across three teacher/student dyads in a high school setting. Overall, the results of the current study were inconsistent with previous literature, with little to no effects observed for TI data and student outcomes. Results should be considered with extreme caution due to the limitations that impacted the completion of the study, and future researchers should be encouraged to continue investigating the effects of the test-driving intervention.

APPENDIX A – Treatment Integrity Forms

Self-Monitoring

	YES	NO
1.Provided student with necessary materials (timer, monitoring form, instructions)		
2. Recorded student behavior every 3 minutes using teacher form		
3. Completed accuracy check with target student		
4. Provided earned rewards or corrective feedback		

Response Cost

	YES	NO
1. Meet and review expectations		
2. Provided student with necessary materials (tokens)		
3. Removes token when student engages in target behavior		
4. Deliver reward contingent on previously agreed upon criteria		

Modified Check-in/Check-out

	YES	NO
1. Meet and review behavior expectations		
2. Remained vigilant of student's behavior throughout session		
3. Completed DBR at the end of session		
4. Reviewed DBR performance with student and provided reward if met		

Behavior Specific Praise

	YES	NO
1. Deliver BSP		
2. Deliver BSP		
3. Deliver BSP		
4. Deliver BSP		
5. Ignore all instances of minor disruptive behavior		

Precorrections

	YES	NO
1. Meet to provide behavioral expectations		
2. Reminds student at 7 or 8 minutes of behavioral expectations		
3. Provides praise for engaging in appropriate behaviors		
4. Provides redirection when engaging in disruptive behaviors		

Self Monitoring

Provide the student with a self-monitoring form and MotivAider® set to four-minute fixed intervals. Tell the student that each time the MotivAider® buzzes, they will determine whether they were on or off-task. If the student is on-task, they should place a check mark in the box. If the student is off-task, they should place an "X" in the box.

The teacher will also monitor the student's behavior using a MotivAider® set to the same interval, and a behavior monitoring form. The teacher will also place a check or an X in the box each time the MotivAider® buzzes.

Accuracy check: At the end of the 15-minute period, the teacher will conduct a check with the student, in which the teacher will determine whether the students monitoring form is accurate. If the students monitoring form matches 80% with the teachers, and they have four or more check marks, they will be provided with a reward.

Student: Date: Teacher:

Place a **check mark** in the box if you were **on-task** when the MotivAider®

Buzzed

Place a **"X"** in the box if you were **off-task** when the MotivAider® Buzzed

1	2	3	4	5

% Agreement with teacher:

Response Cost

Provide the student with five sticky notes. Tell the student that each time they engage in the disruptive behavior, the teacher will take a sticky note away.

During the 15-minute session, take a sticky note away immediately after you observe the student engage in the disruptive behavior. If the student has at least four sticky notes, they may receive access to a reward.

Modified Check-in/Check-out

Meet with the student to review behavior expectations and let them know they must remain "on-task" in the classroom for the majority of the session to earn a reward.

Remain aware of the student's behavior during the next 15 minutes. At the end of the 15-minute session, rate the students on-task behavior on a scale of 0-100% using the Direct Behavior Rating (DBR) form.

After completing the DBR, review the rating with the student, and either provide them with praise and a reward for meeting their goal, or corrective feedback and encouragement to try again later.

Student: Teacher: Date:
Rate how often the student was on-task and circle the appropriate percentage
0%   10%   20%   30%   40%   50%   60%   70%   80%   90%   100%

Behavior Specific Praise

Provide a statement of behavior specific praise (BSP) to the target student at least once every three minutes (or four times in the 15-minute period). Statements of BSP must include a general praise statement such as "good job" or "awesome job" paired with a behavior "being on-task," or "completing your work. Examples of BSP include "Great job staying in your seat," and "Thank you for staying focused on your work" All instances of the student engaging in minor disruptive behavior should be ignored.

Precorrection

Meet with the student to briefly review behavior expectations (e.g., "remember to sit in your seat quietly and complete the assigned worksheet.") After approximately 7-8 minutes, briefly meet with the student to remind them of the expectations.

If the student engages in disruptive behavior, provide a redirection to the current task (e.g., "you need to turn around and face the front and complete your work.") If the student is meeting behavioral expectations, provide praise (e.g., "great job following the instructions."

ORI | Office *of* Research Integrity

INSTITUTIONAL REVIEW BOARD
MINOR ASSENT FORM

MINOR ASSENT PROCEDURES

This document must be completed by the Principal Investigator and signed by each assenting minor.

- The Project Information and Research Description sections of this form should be completed by the Principal Investigator before submitting this form for IRB approval.
- Parental consent must be obtained before soliciting the assent of any minor participating in the study.
- Signed copies of the IRB approved assent form should be provided to a parent or guardian of every assenting minor.

Last Edited March 5th. 2019

Today's date:

PROJECT INFORMATION

Project Title: Test-Driving Interventions for Teachers: Incorporating Treatment Acceptability to Improve Treatment Integrity

Principal Investigator: Stefanie Schrieber	Phone: 309-202-4618	Email: stefschrieb@me.com
College: University of Southern Mississippi	School and Program: School of Psychology // School Psychology	

RESEARCH DESCRIPTION

1. Why am I being asked to participate?

We are looking for interventions that will help you improve your behavior in the classroom. We want to see whether your teachers like implementing any specific interventions and then see if they help you improve your behavior.

2. What will I have to do?

You will just have to listen to your teacher describe the interventions that he/she will be implementing. Your teacher will always provide you with a goal that you will need to meet for your behavior in order to earn a reward. There will not be any extra work that you will have to do other than meeting very briefly with your teacher to determine your goal and whether you met that goal at the end. These interventions will be very short -- only 20 minutes long! You will get to try 4 different interventions, and then your teacher will pick one to implement for about 5-10 days.

3. What do I get if I agree to participate?

If you agree to participate, then each day that you meet your behavior goal, you will get to choose a reward that you would like. This may include free time, a bellwork pass, candy, or chips.

4. Can anything bad happen if I participate?

There are no major risks for participating. You may feel like you are the only one in the classroom receiving the interventions, but there are other students that will be receiving other services. additonally all meetings with your teacher will be confidential and done in private, and no other students will be aware that you are participating.

5. Who will get to see information about me?

All information that we collect will be confidential. Therefore, any data we collect will be de-identified, and no other individuals will know that this is your information. If the researcher publishes this information or presents it at a conference, all names will be changed, and you will not be able to be identified.

6. What if I do not want to participate?

If you decide that you do not want to participate, then we can discuss other behavioral strategies for improving classroom behavior, or we may try a manualized treatment option.

7. Who may I contact if I have other questions or concerns about my participation?

This project has been approved by the Institutional Review Board. Its job is to protect research participants. Questions or concerns about your participation should be directed to the Manager of the IRB at 601-266-5997.

ASSENT TO PARTICIPATE IN RESEARCH	
Participant's Name:	Participant's Age:
Person Soliciting Assent:	

Check one of the following (to be completed by the person soliciting assent):

☐ In my opinion this minor is able to provide informed assent (proceed to Agreement to Participate).
☐ In my opinion this minor is unable to provide informed assent for the following reason(s) (do not proceed):

AGREEMENT TO PARTICIPATE

I agree to participate in this research project. The project has been fully explained to me and I was given the chance to ask any questions I have about it. I understand that I can stop participating at any time.

_____   _____

Research Participant   **Person Soliciting Assent**

_____   _____

Date   **Date**

ORI | Office *of* Research Integrity

INSTITUTIONAL REVIEW BOARD
PARENTAL CONSENT FORM

PARENTAL CONSENT PROCEDURES

This document must be completed by the Principal Investigator and signed by the parent or guardian of each potential research participant.

- The Project Information and Research Description sections of this form should be completed by the Principal Investigator before submitting this form for IRB approval.
- Signed copies of the long form consent should be provided to a parent or guardian of every participant.

Last Edited March 5th, 2019

Today's date:

PROJECT INFORMATION

Project Title: Test-Driving Interventions for Teachers: Incorporating Treatment Acceptability to Improve Treatment Integrity

Principal Investigator: Stefanie Schrieber	Phone: 309-202-4618	Email: stefanie.schrieber@usm.edu
College: University of Southern Mississpi	School and Program: School of Psychology // School Psychology	

RESEARCH DESCRIPTION

1. Purpose:

The purpose of this study is to investigate the effects of an intervention titled Test Driving Interventions for Teachers to increase the integrity with which teachers implement interventions in the classroom. Additionally, this study will provide information about whether improved levels of integrity are related to students academically engaged and disruptive behavior. The results of this study will provide information to other researchers, consultants, and school psychologists to improve other teachers implementation of classroom-based interventions.

2. Description of Study:

Four teacher-student dyads will be recruited based on teacher referral for student problem behavior made to the school's external consultant. Typical behavioral consultation procedures will be used at this point. If the teacher is demonstrationg low treatment integrity (e.g. they are not implementing the intervention as it was intended to be), they will be asked to participate. Student assent will be obtained at this time. The teacher will then implement multiple different evidence-based interventions (20 minutes in duration) with the student across multiple days to determine which intervention they find the most acceptable to implement. The teacher will then implement the most acceptable intervention consistently with the student for approximately 5-10 days. These procedures are consistent with typical behavioral consultation procedures that are conducted at the school. The student will not be responsible for any additional work, nor will they be singled out in class. The selected interventions require a short teacher-student conference (30 seconds) at the beginning of the intervention and at the end.

3. Benefits:

Potential benefits for the student include increased on-task behavior and decreased disruptive behavior. This project may also result in an additional strategy to use to improve teachers intervention implementation

accuracy.

4. Risks:

There are a few risks associated with participation. The student may not be comfortable with the teacher observing their behavior more closely. Additionally, the student may feel singled out, as they may be the only student receiving a targeted intervention in the classroom. These risks will be minimized, as the primary investigator will also be available to answer any questions the student may have. Throughout the experiment, your students' behavior will be monitored. In the event that undesired and unanticipated effects arise (i.e., increase in disruptive behaviors), modifications or termination of procedures will occur and you and your students will be provided with other services. Any potential for students to be singled out will be minimized, as all procedures will be done in confidentiality, and other students will not be informed of the interventions. Additonally, typical consultation procedures will be ongoing for other students present in the classroom.

5. Confidentiality:

All interviews, observations, and other information obtained during this study will be kept strictly confidential. Your name, students' names, and other identifying information will not be disclosed to any person not connected with this study. Results from this research project may be shared at professional conferences or published in scholarly journals; however, all identifying information will be removed from publications and/or presentations.

6. Alternative Procedures:

Should you decide not to participate in the research study, additional strategies will be provided to the teacher to improve student behavior. Additionally, behavioral strategies can be provided to the student and parent.

7. Participant's Assurance:

This project has been reviewed by the Institutional Review Board, which ensures that research projects involving human subjects follow federal regulations.

Any questions or concerns about rights as a research participant should be directed to the Manager of the IRB at 601-266-5997.Participation in this project is completely voluntary, and participants may withdraw from this study at any time without penalty, prejudice, or loss of benefits.

Any questions about the research should be directed to the Principal Investigator using the contact information provided in Project Information Section above.

PARENTAL CONSENT INFORMATION	
Participant's Name:	Participant's Age:
Parent or Guardian's Name:	
Person Soliciting Parental Consent:	

AGREEMENT TO ALLOW PARTICIPATION IN RESEARCH

Consent is hereby given to participate in this research project. All procedures and/or investigations to be followed and their purpose, including any experimental procedures, were explained. Information was given about all benefits, risks, inconveniences, or discomforts that might be expected.

The opportunity to ask questions regarding the research and procedures was given. Participation in the project is completely voluntary, and participants may withdraw at any time without penalty, prejudice, or loss of benefits. All personal information is strictly confidential, and no names will be disclosed. Any new information that develops during the project will be provided if that information may affect the willingness to continue participation in the project.

Questions concerning the research, at any time during or after the project, should be directed to the Principal

Investigator with the contact information provided above. This project and this consent form have been reviewed by the Institutional Review Board, which ensures that research projects involving human subjects follow federal regulations. Any questions or concerns about rights as a research participant should be directed to the Chair of the Institutional Review Board, The University of Southern Mississippi, 118 College Drive #5125, Hattiesburg, MS 39406-0001, 601-266-5997.

Include the following information only if applicable. Otherwise delete this entire paragraph before submitting for IRB approval: The University of Southern Mississippi has no mechanism to provide compensation for participants who may incur injuries as a result of participation in research projects. However, efforts will be made to make available the facilities and professional skills at the University. Participants may incur charges as a result of treatment related to research injuries. Information regarding treatment or the absence of treatment has been given above.

_____   _____
Parent or Guardian of Research Participant   **Person Explaining the Study**

_____   _____
Date   **Date**

Today's date:

PROJECT INFORMATION

| Project Title: Test-Driving Interventions for Teachers: Incorporating Treatment Acceptability to Improve Treatment Integrity |

Principal Investigator: Stefanie Schrieber	Phone: 309-202-4618	Email: stefanie.schrieber@usm.edu
College: University of Southern Mississpi	School and Program: School of Psychology // School Psychology	

RESEARCH DESCRIPTION

1. **Purpose**:

   The purpose of this study is to investigate the effects of an intervention titled Test Driving Interventions for Teachers to determine what interventions teachers find most acceptable. Additionally, this study will provide information about whether different treatments are related to students academically engaged and disruptive behavior. The results of this study will provide information to other researchers, consultants, and school psychologists to improve other teachers implementation of classroom-based interventions.

2. **Description of Study:**

   Four teacher-student dyads will be recruited based on teacher referral for student problem behavior made to the school's external consultant. Typical behavioral consultation procedures will be used at this point and student assent will be obtained at this time. The teacher will then implement multiple different evidence-based interventions (20 minutes in duration) with the student across multiple days to determine which intervention they find the most acceptable to implement. The teacher will then implement the most acceptable intervention consistently with the student for approximately 5-10 days. These procedures are consistent with typical behavioral consultation procedures that are conducted at the school. The student will not be responsible for any additional work, nor will they be singled out in class. The selected interventions require a short teacher-student conference (30 seconds) at the beginning of the intervention and at the end.

3. **Benefits:**

   Potential benefits for the teacher include improvement in student on-task behavior and decreased disruptive behavior. This project may also result in an additional strategy to use to improve teachers intervention implementation accuracy.

4. **Risks:**

There are a few risks associated with participation. The student may not be comfortable with the teacher observing their behavior more closely. Additionally, the student may feel singled out, as they may be the only student receiving a targeted intervention in the classroom. These risks will be minimized, as the primary investigator will also be available to answer any questions the student may have. Throughout the experiment, your students' behavior will be monitored. In the event that undesired and unanticipated effects arise (i.e., increase in disruptive behaviors), modifications or termination of procedures will occur and you and your students will be provided with other services. Any potential for students to be singled out will be minimized, as all procedures will be done in confidentiality, and other students will not be informed of the interventions. Additonally, typical consultation procedures will be ongoing for other students present in the classroom.

5. **Confidentiality:**

All interviews, observations, and other information obtained during this study will be kept strictly confidential. Your name, students' names, and other identifying information will not be disclosed to any person not connected with this study. Results from this research project may be shared at professional conferences or published in scholarly journals; however, all identifying information will be removed from publications and/or presentations.

6. **Alternative Procedures:**

Should you decide not to participate in the research study, additional strategies will be provided to the teacher to improve student behavior. Additionally, behavioral strategies can be provided to the student and parent.

7. **Participant's Assurance:**

This project has been reviewed by the Institutional Review Board, which ensures that research projects involving human subjects follow federal regulations.

Any questions or concerns about rights as a research participant should be directed to the Manager of the IRB at 601-266-5997.Participation in this project is completely voluntary, and participants may withdraw from this study at any time without penalty, prejudice, or loss of benefits.

Any questions about the research should be directed to the Principal Investigator using the contact information provided in Project Information Section above.

regulations. Any questions or concerns about rights as a research participant should be directed to the Chair of the Institutional Review Board, The University of Southern Mississippi, 118 College Drive #5125, Hattiesburg, MS 39406-0001, 601-266-5997.

_____   _____

Teacher   Person Explaining Study

_____   _____

Date   Date

APPENDIX F – IRB Approval Letter

Modification Institutional Review Board Approval

The University of Southern Mississippi's Office of Research Integrity has received the notice of your modification for your submission Test-Driving Interventions for Teachers: Incorporating Treatment Acceptability to Improve Treatment Integrity (IRB #: IRB-19-373).

Your modification has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services regulations (45 CFR Part 46), and University Policy to ensure:

- The risks to subjects are minimized and reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered involving risks to subjects must be reported immediately. Problems should be reported to ORI via the Incident template on Cayuse IRB.
- The period of approval is twelve months. An application for renewal must be submitted for projects exceeding twelve months.

PROTOCOL NUMBER: IRB-19-373
PROJECT TITLE: Test-Driving Interventions for Teachers: Incorporating Treatment Acceptability to Improve Treatment Integrity
SCHOOL/PROGRAM: School of Psychology, Psychology
RESEARCHER(S): Stefanie Schrieber ,TERM Evan TERM Dart, Brad Dufrene, Taylor Ben, Mary Ware

IRB COMMITTEE ACTION: Approved
7. Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

PERIOD OF APPROVAL: February 10, 2020

Donald Sacco, Ph.D.
Institutional Review Board Chairperson

APPENDIX G – Tables and Figures

Table A1. *Mean Teacher Ratings on the URP-IR (Adapted)*

	Response Cost	BSP	CICO	Precorrections	Self-Monitoring
Dyad 1 (Caroline)					
Acceptability	**5.11**	4.67	3.67	3.56	4.89*
Feasibility	**5.17**	5.50	5.50	5.33	4.83*
Dyad 2 (Lexi)					
Acceptability	5.67*	5.78	4.67	4.67	**5.78**
Feasibility	5.50*	5.83	4.17	5.50	**5.83**
Dyad 3 (Bonnie)					
Acceptability			4.89*		
Feasibility			4.50*		

(Bold) Denotes teacher's most preferred intervention, (*) denotes consultation-as-usual intervention

Table A2. *Mean Teacher Ratings on the URP-A*

	Dyad 1 (Caroline)	Dyad 2 (Lexi)
Acceptability	4.78	5.89
Understanding	5.67	6
Home/school Collaboration	3.67	5
Feasibility	5	5.67
System Climate	5.25	6
System Support	4	1.33

Figure 1. Treatment Integrity and Student Behavior Data
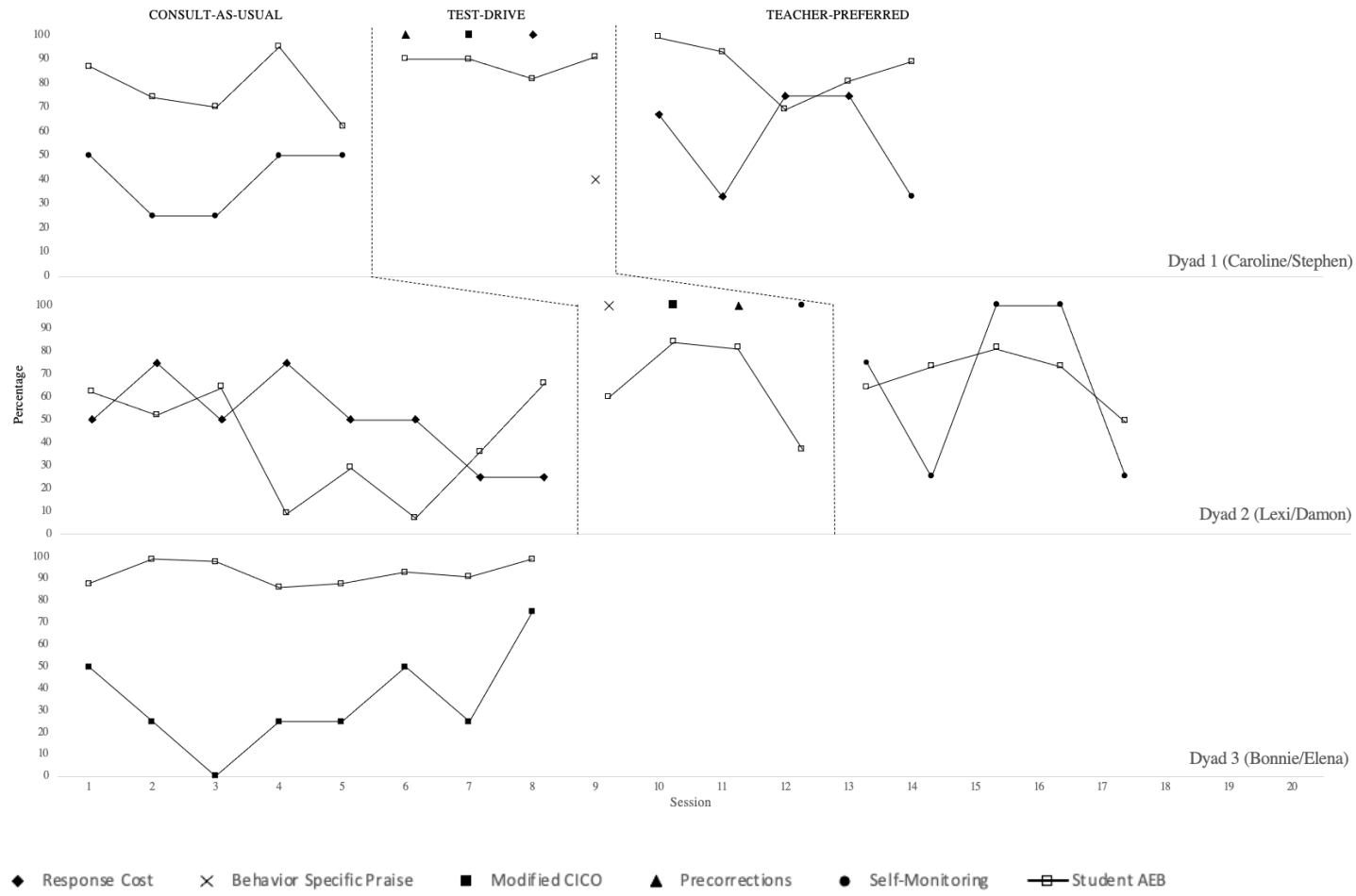
REFERENCES

Allinder, R. M., & Oats, R. G. (1997). Effects of acceptability on teachers'

    implementation of curriculum-based measurement and student achievement in

    mathematics computation. *Remedial and Special Education, 18,* 113–120.

Amato-Zech, N. A., Hoff, K. E., Doepke, K. J. (2006). Increasing on-task behavior in the

    classroom: Extension of self-monitoring strategies. *Psychology in the Schools, 43,*

    211-231.

Barrish, H., Saunders, M., & Wolf, M. (1969). Good Behavior Game: Effects of

    individual contingencies for group consequences on disruptive behavior in a

    classroom. *Journal of Applied Behavior Analysis*, *2*, 119-124.

    doi:10.1901/jaba.1969.2-119

Bergan, J. R., & Kratochwill, T. R. (1990). *Behavioral consultation and therapy.* New

    York,   NY: Plenum.

Biggs, B. K., Vernberg, E. M., Twemlow, S. W., Fonagy, P., & Dill, E. J. (2008).

    Teacher adherence and its relation to teacher attitudes and student outcomes in an

    elementary school-based violence prevention program. *School Psychology*

    Review, 37, 533-549.

Briesch, A. M., Chafouleas, S. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2013).

    Assessing influences on intervention use: Revision of the Usage Rating

    Profile-Intervention. *Journal of School Psychology, 51*, 81–96.

Brown, S., & Rahn-Blakeslee, A. (2009). Training school-based practitioners to collect

    intervention integrity data. *School Mental Health, 1*, 143–153.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd
ed.). Upper Saddle River, NJ: Pearson.

Dart, E. H., Cook, C. R., Collins, T. A., Gresham, F. M., & Chenier, J. S. (2012). Test-
driving interventions to increase treatment integrity and student outcomes. School
Psychology Review, 41, 467– 481.

DiGennaro, F. D., Martens, B. K., & Kleinmann, A. E. (2007). A comparison of
performance feedback procedures on teachers' treatment implementation integrity
and students' inappropriate behavior in special education classrooms. Journal of
Applied Behavior Analysis, 40, 447-461.

Erchul, W. P., & Martens, B. K. (2012). School consultation: Conceptual and empirical
bases of practice: 3rd ed. New York: Springer.

Fallon, L. M., Collier-Meek, M. A., Maggin, D. M., Sanetti, L. M. H., & Johnson, A. H.
(2015). Is performance feedback an evidence-based intervention? A systematic
review and evaluation of single-case research. Exceptional Children, 8, 227–246.

Fallon, L. M., Kurtz, K. D., & Mueller, M. R. (2018). Direct training to improve
educators' treatment integrity: A systematic review of single-case design studies.
School Psychology Quarterly, 33(2), 169–181.

Forman, S. G., Shapiro, E. S., Codding, R. S., Gonzales, J. E., Reddy, L. A., Rosenfield,
S. A., Sanetti, L. M. H., & Stoiber, K. C. (2013). Implementation science and
school psychology. *School Psychology Quarterly, 28*, 77–100.

Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and
prereferral intervention. *School Psychology Review, 18*, 37–50.

Gresham, F. M. (2005). Response to intervention: An alternative means to identifying

students as emotionally disturbed. *Education and Treatment of Children, 28*, 328–

344.

Gresham, F. M. (2010). Data-based decision making for students' social behavior.

Journal of Evidence-Based Practices for Schools, 11, 149-168.

Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M.

(2000). Treatment integrity in learning dis- abilities intervention research: Do we

really know how treatments are implemented? *Learning Disabilities Research*

and Practice, 15, 198–205.

Gresham, F. M., Dart, E. H., & Collins, T. A. (2017). Generalizability of multiple

measures of treatment integrity: Comparisons among direct observation,

permanent products, and self-report. *School Psychology Review, 46,* 108-121.

Gutkin, T. B., & Conoley, J. C. (1990). Reconceptualizing school psychology from a

service delivery perspective: Implications for practice, training, and research.

Journal of School Psychology, 28, 203-223.

Gutkin, T. B. (1996). Patterns of consultant and consultee verbalizations: Examining

communication leadership during initial consultation interviews. Journal of

School Psychology, 34, 199-219.

Hagermoser, Sanetti, L., & Kratochwill, T. R. (2005). Treatment integrity assessment

within a problem-solving model. In R. Brown-Chidsey (Ed.), Assessment for

intervention: A problem-solving approach (pp.304-325). New York: Guilford

Press.

Hawken, L. H., & Horner, R. H. (2003). Evaluation of a targeted intervention within a

    schoolwide system of behavior support. *Journal of Behavioral Education, 12,*

    225–240.

Hettema, J., Ernst, D., Williams, J. R., & Miller, K. J. (2014). Parallel processes: Using

    motivational interviewing as an implementation coaching strategy. *Journal of*

    Behavioral Health Services and Research, 41, 324–336.

Jones, K. M., Drew, H. A., & Weber, N. L. (2000). Noncontingent peer attention as

    treatment for disruptive classroom behavior. *Journal of applied behavior analysis,*

    33, 343-346.

Kazdin, A. E. (1980). Acceptability of alternative treatments for deviant child behavior.

    Journal of Applied Behavior Analysis, *13,* 259–273.

Kilgus, S. P., Riley-Tillman, T. C., Chafouleas, S. M., Christ, T. J., & Welsh, M. E.

    (2014). Direct behavior rating as a school-based behavior universal screener:

    Replication across sites. *Journal of School Psychology, 52,* 63-82.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D.

    M., & Shadish, W. R. (2010). *Single-case designs technical documentation*.

    Retrieved from What Works Clearinghouse web site: http://

    ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Kratochwill, T. R., & Stoiber, K. C. (2000). Empirically supported interventions and

    school psychology: Rationale and methodological issues—Part I. *School*

    Psychology Quarterly, 15, 75-105.

Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment

 integrity: An essential but often forgotten component of school-based

 interventions. *Preventing School Failure, 48,* 36–43.

McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of

 evidence-based treatments. *American Psychologist, 65*, 73–84.

Miller WR. 1983. Motivational interviewing with problem drinkers. *Behavioral*

 Psychotherapy, 11, 147–172.

Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change*

 (3rd ed.). New York: Guilford Press.

Mortenson, B. P., & Witt, J. C. (1998). The use of weekly performance feedback to

 increase teacher implementation of a prereferral academic intervention. *School*

 Psychology Review, 27(4), 613-627.

MotivAider (2017). *MotivAider®: Making people better, one habit at a time.*

 (Behavioral Dynamics) Retrieved from habitchange.com

Musser, E. H., Bray, M. A., Kehle, T. J., & Jenson, W. R. (2001). Reducing disruptive

 behaviors in students with serious emotional disturbance. *School Psychology*

 Review, 30, 294–304.

Neugebauer, S. R., Chafouleas, S. M., Coyne, M. D., McCoach, D. B., Briesch, A. M.

 (2016). Exploring an ecological model of perceived usability within a multi-tiered

 vocabulary intervention. *Assessment for Effective Intervention, 41,* 155-171.

Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997).

 Increasing teacher intervention implementation in general education settings

through consultation and performance feedback. *School Psychology Quarterly, 12*, 77-78.

Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., Koenig, J. L., & Resetar, J. L. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. School Psychology Review, 34, 87-106.

Noell, G. H. (2010). Empirical and pragmatic issues in assessing and supporting intervention implementation in school. In G. G. Peacock, R. A. Ervin, E. J. Daly, & K. W. Merrell (Eds.), *Practical handbook in school psychology* (pp. 513–530). New York, NY: Guilford.

Al Otaiba, S., Wagner, R. K., Miller, B. (2014). "Waiting to fail" redux: Understanding inadequate response to intervention. *Learning Disabilities Quarterly, 37,* 129-133.

Peterson, L., Homer, A., & Wonderlich, S. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis, 15*, 477–492.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: The Free Press.

Sanetti, L. M. H., & Collier-Meek, S. M. (2015). Data-driven delivery of implementation supports in a multi-tiered framework: A pilot study. *Psychology in the Schools, 52*, 815–828.

Sanetti, L. M. H., & Kratochwill, T. R. (2008). Treatment integrity in behavioral consultation: Measurement, promotion, and out- comes. *International Journal of Behavioral Consultation and Therapy, 4*, 95–114.

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to a special series. *School Psychology Review, 38*, 445–459.

Sanetti, L. M. H., & Kratochwill, T. R. (2011). Daily versus weekly teacher self-report of treatment integrity: Impact on implementation and report accuracy. *Journal of Educational and Psychological Consultation*, *21*, 284-308.

Sanetti, L. M. H., Collier-Meek, M. A., Long, A. C. J., Byron, J. R., & Kratochwill, T. R. (2015). Increasing teacher treatment integrity of behavior support plans through consultation and implementation planning. *Journal of School Psychology, 53*, 209–229.

Sanetti, L. M. H., Fallon, L. M., Collier-Meek, M. A. (2011). Treatment integrity assessment and intervention by school-based personnel: Practical applications based on a preliminary study. *School Psychology Forum, 5*, 87–102.

Sanetti, L. M. H., Gritter, K. L., & Dobey, L. M. (2011). Treatment integrity of interventions with children in the school psychology literature from 1995 to 2008. *School Psychology Review*, *40*, 72-84.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin.

Sheridan, S. M., Welch, M., & Orme, S. F. (1996). Is consultation effective? A review of outcome research. *Remedial and Special Education, 17,* 341–354.

Silva, M. R., Collier-Meek, M. A., Codding, R. S., & Defouw, E. R. (2019) Acceptability assessment of school psychology interventions from 2005 to 2017. *Psychology in the Schools, 1-16.*

Simonsen, B., MacSuga, A. S., Fallon, L. M., & Sugai, G. (2013). The effects of self-monitoring on teachers' use of specific praise. *Journal of Positive Behavior Interventions*, *15*, 5–15.

Simonsen, B., Myers, D., & DeLuca, C. (2010). Providing teachers with training and performance feedback to increase use of three classroom management skills: Prompts, opportunities to respond, and reinforcement. *Teacher Education in Special Education*, *33*, 300–318.

Solomon, B. G., Klein, S. A., & Politylo, B. C. (2012). The effect of performance feedback on teachers' treatment integrity: A meta-analysis of the single-case literature. *School Psychology Review, 41*, 160–175.

Sterling-Turner, H. E., Watson, T. S. (2002). An analog investigation of the relationship between treatment acceptability and treatment integrity. *Journal of Behavioral Education, 11*, 39- 50.

Sterling-Turner, H. E., Watson, T. S., Wildmon, M., Watkins, C., & Little, E. (2001). Investigating the relationship between training type and treatment integrity. *School Psychology Quarterly, 16,* 78–89.

Thompson, R. H., & Iwata, B. A. (2005). A review of reinforcement control procedures. *Journal of Applied Behavior Analysis, 38,* 257-278.

Tiano, J. D., Fortson, B. L., McNeil, C. B., & Humphreys, L. A. (2005). Managing classroom behavior of Head Start children using response cost and token economy procedures. *Journal of Early and Intensive Behavior Intervention, 2*(1), 28-39.

Tiger, J. H., Hanley, G. P., & Hernandez, E. (2006). *An Evaluation of the Value of Choice with Preschool Children. Journal of Applied Behavior Analysis, 39(1),* 1-16.

Todd, A. W., Campbell, A. L., Meyer, G. G., & Horner, R. H. (2008). The effects of a targeted intervention to reduce problem behaviors: Elementary school implementation of check in–check out. *Journal of Positive Behavior Interventions, 10,* 46–55.

Wenger, R. (1979). Teacher response to collaborative consultation. *Psychology in the Schools, 16,* 127-131.

Wickstrom, K. F., Jones, K. J., LaFleur, L. H., & Witt, J. C. (1998). An analysis of treatment integrity in school-based behavioral consultation. *School Psychology Quarterly, 13,* 141– 154.

Wilkinson, L. A. (2007). Assessing treatment integrity in behavioral consultation. *International   Journal of Behavioral Consultation and Therapy, 3,* 420-432.

Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), Advances in school psychology (Vol. 6, pp. 252-288). Hillsdale, NJ: Erlbaum.